



(19)

Europäisches Patentamt

European Patent Office

Office pour le dépôt des brevets



(11)

EP 0 745 675 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
04.12.1996 Bulletin 1996/49

(51) Int. Cl.⁶: C12N 15/54, C12N 9/12,
C12N 1/21, C12N 15/70,
C12Q 1/68
// (C12N1/21, C12R1:19)

(21) Application number: 96108613.9

(22) Date of filing: 30.05.1996

(84) Designated Contracting States:
CH DE FR GB IT LI

(30) Priority: 31.05.1995 JP 134096/95

(71) Applicant: Toyo Boseki Kabushiki Kaisha
Osaka-shi Osaka 530 (JP)

(72) Inventors:

- Kitabayashi, Masao,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)
- Arakawa, Taku,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)
- Inoue, Hiroaki,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)

- Kawakami, Bunsei,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)
- Kawamura, Yoshihisa,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)
- Imanaka, Takayuki
Suita-shi, Osaka-fu (JP)
- Takagi, Masahiro
Suita-shi, Osaka-fu (JP)
- Morikawa, Masaaki
Minoo-shi, Osaka-fu (JP)

(74) Representative: VOSSIUS & PARTNER
Postfach 86 07 67
81634 München (DE)

(54) A method of amplifying nucleic acid and a reagent therefor

(57) An object of this invention is to provide an enzyme which amplifies nucleic acid with a high fidelity within a short reaction time and also to provide a method of amplification.

This invention relates to a thermostable DNA polymerase having at least 30 bases/second of DNA extension rate and a 3'-5' exonuclease activity derived from an hyperthermophilic archaeon strain KOD1, to a method of amplifying and also of detecting the nucleic acid using said enzyme and to a reagent kit used for those methods.

Description

The present invention relates to a method of amplifying nucleic acid wherein DNA or RNA is amplified within a short reaction time and with a high fidelity, to a method of identifying nucleic acid utilizing said amplifying method and to a 5 DNA polymerase and a reagent kit used for those methods.

Many studies have been made already for DNA polymerase of mesophilic microorganisms such as Escherichia coli and for DNA polymerase derived from phages infectable by the mesophilic microorganisms. In addition, many studies have been also made already for heat stable DNA polymerases which are useful in a recombinant DNA technique by 10 means of nucleic acid amplification such as a polymerase chain reaction (PCR). Examples of the heat-stable polymerases which are used for the PCR are DNA polymerase (Tth polymerase) mostly derived from Thermus thermophilus and DNA polymerase (Taq polymerase) derived from Thermus aquaticus. Other known examples are DNA polymerase (Pfu polymerase) derived from Pyrococcus furiosus and DNA polymerase (Vent polymerase) derived from Thermococcus litoralis.

However, with the Taq polymerase, fidelity and thermostability upon the synthesis of DNA are not sufficient.

15 Although the Pfu polymerase exhibiting excellent fidelity and thermostability has been developed, said Pfu polymerase has some problems that its DNA extension rate is slow and a processivity is low whereby it has been used only for a specific PCR.

Recently, a PCR whereby 20 kb or more DNA is amplified (hereinafter, referred to as a long-PCR) has been developed. In said long-PCR, both Taq polymerase and Pfu polymerase are mixed whereby properties of both enzymes are 20 utilized. However, when two enzymes having different properties are used in the same reaction system, some discrepancies might occur in their appropriate reaction conditions whereby there is a question whether the high extension rate and fidelity which are the advantages of each of those enzymes can be still maintained. Moreover, because of the difference in the thermostabilities and in the composition of the stock solutions of both enzymes, there is a question as to the stability when they are stoked in the same container.

25 In view of the above, there has been a keen demand for novel thermostable polymerase which exhibits both of those advantages.

The present inventors have succeeded in preparing a thermostable DNA polymerase from a hyperthermophilic archaeon strain KOD1, and, when its properties are investigated, it has been found that said DNA polymerase exhibits the advantages of the above-mentioned two enzymes, i.e. high extension rate and high fidelity, whereby the present 30 invention has been achieved.

Thus, the present invention relates to a method for amplifying a target nucleic acid comprises reacting the target nucleic acid with four kinds of dNTP and primer complementary to said target nucleic acid in a buffer solution which contains a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity such that the above mentioned primer is annealed to the target nucleic acid and an extention product is synthesized from the primer.

The present invention further relates to a method for amplifying a target nucleic acid in a sample wherein each target nucleic acid consists of two separate complementary strands which comprises the following steps A to D, characterized in that a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity is used as a thermostable DNA polymerase;

40 A: modifying the target nucleic acid, if necessary, to produce single-stranded nucleic acids;

B: reacting the single-stranded nucleic acids with four kinds of dNTP and primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, in a buffer solution which contains a thermos table DNA polymerase such that the above mentioned primers are annealed to the Single-stranded nucleic acids and extention products are synthesized from the primers,

45 C: separating the primer extention products from the templates on which they are synthesized to produce single-stranded nucleic acids; and

D: repeatedly conducting the above mentioned steps B and C.

50 The present invention further relates to a method for detecting a target nucleic acid in a sample wherein each target nucleic acid consists of two separate complementary strands which comprises the following steps A to E, characterized in that a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity is used as a thermostable DNA polymerase;

55 A: modifying the target nucleic acid, if necessary, to produce single-stranded nucleic acids;

B: reacting the single-stranded nucleic acids with four kinds of dNTP and primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, in a buffer solution which contains a thermostable DNA polymerase such that the above mentioned primers are annealed to the single-stranded nucleic acids and extention products are synthesized from the primers,

C: separating the primer extention products from the templates on which they are synthesized to produce single-stranded nucleic acids;

D: repeatedly conducting the above mentioned steps B and C, and

E: detecting an amplified nucleic acid.

5 The present invention further relates to a reagent kit for amplifying target nucleic acid which comprises primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, four kinds of dNTP, divalent cation, thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity and buffer solution.

10 The present invention further relates to a reagent kit for detecting target nucleic acid which comprises primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, four kinds of dNTP, divalent cation, thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity, amplifying buffer solution, a probe capable of hybridizing with amplified nucleic acid and a detection buffer solution.

15 The present invention relates to a thermos table DNA polymerase which is obtainable from a strain KOD1 which belongs to a hyperthermophilic archaeon strain.

The present invention relates to an isolated DNA comprising a nucleotide sequence that encodes the thermostable DNA polymerase derived from a KOD1 strain which belongs to hyperthermophilic archaeon.

20 Preferably, said DNA comprises the nucleotide sequence of SEQ ID No. 5 or a nucleotide sequence encoding the thermostable polymerase which has the amino acid sequence of SEQ ID No. 1.

The present invention also relates to a DNA encoding a polypeptide with the above mentioned biological activities comprising a nucleotide sequence

(a) which differs from the above DNA in codon sequence due to the degeneracy of the genetic code;

25 (b) which hybridizes with the above DNA or the DNA of section (a); or

(c) represents a fragment, allelic or other variation of the above DNA.

30 In this context, the term "hybridization" refers to conventional hybridization conditions (see for example Nucleic acid hybridization, A practical approach, Hames and Higgins, eds., IRL Press, Oxford Washington DC, 1985). Most preferably, the term "hybridization" refers to stringent hybridization conditions.

The present invention also relates to a thermostable DNA polymerase with the above mentioned biological activities and being encoded by one of said DNA's and, in addition, to a kit comprising said thermostable DNA polymerase.

35 The present invention further relates to a recombinant DNA expression vector that comprises the DNA sequence inserted into a vector, wherein the DNA sequence encodes the thermostable DNA polymerase derived from a KOD1 strain which belongs to hyperthermophilic archaeon.

The present invention further relates to a transformed recombinant host cell using a recombinant DNA expression vector that comprises the DNA sequence inserted into a vector, wherein the DNA sequence encodes the thermos table DNA polymerase derived from a KOD1 strain which belongs to hyperthermophilic archaeon.

40 The present invention relates to a method for producing a DNA polymerase obtainable from a KOD1 strain which belongs to hyperthermophilic archaeon, comprises culturing recombinant host cells which is transformed by a recombinant DNA expression vector that comprises the DNA sequence inserted into a vector, wherein the DNA sequence encodes the thermostable DNA polymerase derived from a KOD1 strain which belongs to hyperthermophilic archaeon, and recovering the produced thermostable DNA polymerase.

45 The present invention further relates to a method for purifying the DNA polymerase obtainable from a KOD1 strain which belongs to hyperthermophilic archaeon, comprises culturing the recombinant host cells which is transformed by a recombinant DNA expression vector that comprises the DNA sequence inserted into a vector, wherein the DNA sequence encodes the thermos table DNA polymerase derived from a KOD1 strain which belongs to hyperthermophilic archaeon, and further (a) recovering the cultured recombinant host cells, disintegrating them and preparing the cell extract, and (b) removing the impurified proteins derived from recombinant host cells.

50 The nucleic acid which is to be amplified by the present invention is DNA or RNA. There is no restriction at all for the sample in which such a nucleic acid is contained.

The thermostable enzyme which is used in the present invention is a thermostable DNA polymerase having at least 30 bases/second of DNA extension rate and having a 3'-5' exonuclease activity. Its specific example is a DNA polymerase derived from a hyperthermophilic archaeon strain KOD1 (called a KOD polymerase) and said enzyme may be either a thermostable enzyme purified from nature or an enzyme manufactured by a gene recombination technique.

55 The DNA extension rate in the present invention is calculated from the relationship between the reaction time and the size of the synthesized DNA in the reaction of various kinds of DNA polymerases such as KOD, Pfu, Deep Vent, Taq, etc. (5U) in each buffer using a substrate prepared by annealing a single-stranded DNA (1.6 µg) of M13 with a primer (16 pmoles) complementary thereto. It is essential in the present invention that the DNA extension rate is at least

30 bases/second.

The DNA extension rates for each of the polymerases are 105-130 bases/second for KOD polymerase, 24.8 bases/second for Pfu polymerase, 23.3 bases/second for Deep Vent polymerase and 61.0 bases/second for Taq polymerase.

5 On the other hand, it is essential in the present invention that the thermostable DNA polymerase has a 3'-5' exonuclease activity.

In the present invention, the 3'-5' exonuclease activity is determined by checking the rate of release of ^{3}H under the optimum condition for each polymerase using a substrate wherein the 3'-end of the lambda-DNA digested with HindIII labeled with $[^{3}\text{H}]$ TTP.

10 In the 3'-5' exonuclease activity of each polymerase, free ^{3}H is found to be only 10-20% in the case of Taq polymerase and Tth polymerase after an incubation period of three hours, in KOD polymerase and Pfu polymerase, it is 50-70%.

It has been confirmed that the KOD polymerase used in the Present invention has a 3'-5' exonuclease activity and that, in the gene which codes for KOD polymerase, there is a DNA conserved sequence showing a 3'-5' exonuclease 15 activity the same as in the case of Pfu polymerase.

In the present invention, the fact whether there is a 3'-5' exonuclease activity is checked in such a manner that KOD polymerase is allowed to stand, using a DNA fragment into which the DNA of $[^{3}\text{H}]$ TTP-labelled-lambda-DNA digested with HindIII is incorporated as a substrate, at the reaction temperature of 75°C in a buffer (20mM Tris-HCl of pH 6.5, 10mM KCl, 6mM $(\text{NH}_4)_2\text{SO}_4$, 2mM MgCl₂, 0.1% Triton X-100 and 10 $\mu\text{g}/\text{ml}$ BSA) and the ratio of the free- $[^{3}\text{H}]$ TTP is 20 determined.

At the same time, Taq polymerase and Tth polymerase having no 3'-5' exonuclease activity and Pfu polymerase having a 3'-5' exonuclease activity were checked using a buffer for each of them by the same manner as in the control experiments. The titer of each of the used polymerases was made 2.5 units.

The substrate DNA was prepared in such a manner that, first, 0.2 mM of dATP, dGTP, dCTP and $[^{3}\text{H}]$ TTP were 25 added to 10 μg of lambda-DNA digested with HindIII, the 3'-end was elongated by Klenow polymerase, then DNA fragments were recovered by extracting with phenol and precipitating with ethanol and free mononucleotides were removed by a Spin column (manufactured by Clontech).

In the case of KOD polymerase and Pfu polymerase, 50-70% of free $[^{3}\text{H}]$ TTP were detected after an incubation period of three hours, in the case of Taq polymerase and Tth polymerase, only 10-20% of free $[^{3}\text{H}]$ TTP was noted.

30 It is preferred that said thermostable DNA polymerase contains an amino acid sequence given in SEQ ID No.1.

It is also preferred that said thermostable DNA polymerase is an enzyme having the following physical and chemical properties.

Action: It has a DNA synthetic activity and a 3'-5' exonuclease activity.

DNA extension rate: at least 30 bases/second

35 Optimum pH: 6.5-7.5 (at 75°C)

Optimum temperature: 75°C

Molecular weight: about 88-90 Kda

Amino acid sequence: as mentioned in SEQ ID No.1

An example of the methods for manufacturing DNA polymerase derived from a hyperthermophilic archaeon strain 40 KOD1 is that thermostable DNA polymerase gene was cloned from strain KOD1 which was isolated from a softtara at a wharf on Kodakara Island, Kagoshima so that a recombinant expression vector was constructed, then a transformant prepared by transformation by said recombinant vector was cultured and the thermostable DNA polymerase was collected from the culture followed by purifying.

In the present invention, the DNA polymerase derived from the above-mentioned hyperthermophilic archaeon 45 strain KOD1 has a DNA synthesizing activity and a 3'-5' exonuclease activity and has a DNA extension rate of at least 30 bases/second. This property is used for conducting an amplification of nucleic acid

The amplifying method of the present invention includes the following steps A to D.

50 A: modifying the target nucleic acid, if necessary, to produce single-stranded nucleic acids;

B: reacting the single-stranded nucleic acids with four kinds of dNTP and primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, in a buffer solution which contains a thermostable DNA polymerase such that the above mentioned primers are annealed to the single-stranded nucleic acids and extention products are synthesized from the primers,

55 C: separating the primer extention products from the templates on which they are synthesized to produce single-stranded nucleic acids; and

D: repeatedly conducting the above mentioned steps B and C.

In the step A, the target nucleic acid is denatured if necessary to give a single-stranded nucleic acid. The means therefor may be a thermal treatment, a chemical denaturation or an enzymatic treatment. Preferably, it is a thermal

treatment.

In the step B, said single-stranded nucleic acid is made to react with four kinds of dNTP (dATP, dGTP, dCTP and dTTP or dUTP) and primers with regular and inverted directions having complementary base sequences to the target nucleic acid in a buffer solution containing a thermostable DNA polymerase so that said primers are annealed to the single-stranded nucleic acid to conduct a primer extension reaction.

5 A primer with a regular direction and that with an inverted direction having complementary base sequences to the target nucleic acid are oligonucleotides having a base sequence which is complementary to one of target nucleic acid and is homologous to another. Accordingly, one primer may be complementary to another primer elongate.

Preferred buffer solutions containing a thermostable DNA polymerase are Tris buffers containing divalent cation 10 such as magnesium ion.

An example of the conditions for conducting an elongation reaction by annealing the primer is a method in which a cycle of 98°C/1 second-1 minute and 68°C/1 second-10 minutes is repeated for 30 times.

The step of separating an elongated primer for making a single strand in the step C may be a thermal treatment, a 15 chemical treatment or an enzymatic treatment. Preferably, it is a thermal treatment or an enzymatic treatment using RNase.

In the step D, the above-mentioned steps B and C are repeated. To be more specific, it is preferred that heating and cooling of 98°C/20 seconds and 68°C/30 seconds are repeated at least for 30 cycles.

An amplifying method of the present invention is applicable to a PCR for amplifying a DNA of 20 kb or more (hereinafter, referred to as a long-PCR) as well. In this long-PCR, advantages of both high DNA extension rate of Taq 20 polymerase and high fidelity in DNA synthesis caused by a 3'-5' exonuclease activity of Pfu polymerase are necessary and both enzymes are used after mixing them. In this case, there is a question on a stability when both enzymes are stored in the same container because of the difference between their thermostabilities and that between the compositions of their stored solutions. However, in the DNA polymerase derived from a hyperthermophilic archaeon strain KOD1, a single enzyme exhibits both high DNA extension rate and high fidelity due to its 3',-5' exonuclease activity 25 whereby it is possible that a long-PCR can be conducted by its sole use.

In the present invention, the amplified product produced by the above-mentioned amplification such as a labeled probe is used whereby a target nucleic acid can be detected.

Labeled probe is an oligonucleotide having a base sequence which is complementary to a target nucleic acid and is bonded with a labeled substance or a labeled binding substance.

30 Examples of the labeled substance are enzymes such as alkaline phosphatase, peroxidase and galactosidase, fluorescent substances and radioactive substances while examples of the labeled binding substances are biotin and digoxigenin. Labeled substance may be bonded via biotin, digoxigenin or avidin.

A method of introducing those labels into a probe is that, during the synthesis of oligonucleotide, dNTP to which 35 those labeled substances or labeled binding substances are bonded is used as one of the components of dNTP whereby a synthesis is conducted.

Examples of detecting a nucleic acid bonded with a labeled probe are conventionally known methods such as a Southern hybridization and a Northern hybridization. In those methods, the fact that a hybrid is formed when single-stranded DNA and RNA are complementary each other is utilized whereby unknown nucleic acid fraction group is subjected to an agarose electrophoresis to separate its size, then the nucleic acid fraction in the gel is subjected, for example, to an alkali treatment, the resulting single strand is transferred to a filter, immobilized and hybridized with a labeled probe.

40 As to a detection of the label in case an alkaline phosphatase is used as a labeled substance, when a chemoluminescent substrate such as a 1,2-dioxetane compound (PPD) is made to react therewith, only nucleic acid forming a hybrid is illuminated. This is sensitized to an X-ray film whereby the size of the target nucleic acid and its position on electrophoresis can be confirmed.

A reagent kit for nucleic acid amplification according to the present invention contains primers of regular and inverted directions having base sequences complementary to target nucleic acid, four kinds of dNTP, divalent cation, thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and having a 3'-5' exonuclease activity and a buffer solution.

45 An example of divalent cation is magnesium ion. Its concentration is preferably about 1-3 mM. Examples of the buffer solution are tris buffer (pH 6.5, 75°C) and tricine buffer (pH 6.5, 75°C).

A specific example of the composition is as follows:

20mM Tris-HCl (pH 6.5, 75°C)

10mM KCl

55 6mM (NH₄)₂SO₄

1-3mM MgCl₂

0.1% Triton X-100

10 µg/ml BSA

20-200 µM dNTPs

0.1pM-1μM primer

0.1-250ng template DNA.

A reagent kit for nucleic acid amplification according to the present invention contains a nucleic acid amplifying reagent comprising primers of regular and inverted directions having base sequences complementary to target nucleic acid, four kinds of dNTP, divalent cation, thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and having a 3'-5' exonuclease activity and a buffer solution for amplification, a target nucleic acid probe and a buffer for detection. The buffer for detection is that the detecting reagent varies depending upon the label. For example, it includes a color reagent or a luminous reagent.

5 KOD1 which is a kind of hyperthermophilic archaeon used in the present invention is a strain isolated from a solfatara at a wharf on Kodakara Island, Kagoshima.

10 Mycological properties of said strain are as follows.

Shape of cells: coccus, diplococcus; having flagella.

Temperature range for the growth: 65-100°C

Optimum temperature for the growth: 95°C

15 pH range for the growth: 5-9

Optimum pH: 6

Optimum salt concentration: 2-3%

Auxotrophy: heterotrophic

Oxygen demand: aerophilic

20 Cell membrane lipids: ether type

GC content of DNA: 38%

The hyperthermophilic archaeon strain KOD1 was a coccus having a diameter of about 1 μm and had plural polar flagella. From the mycological properties of the strain, its close relationship with Pfu DNA polymerase-productive bacterium (*Pyrococcus furiosus*) and with Tli (Vent) DNA polymerase-productive bacterium (*Thermococcus litoralis*) was suggested.

25 Cloning of the thermostable DNA polymerase gene of the present invention is carried out as follows.

The cloning method is that a primer is designed and synthesized depending upon an amino acid sequence in a conserved region of Pfu DNA polymerase (Nucleic Acids Research, 1993, vol.21, No.2, 259-265).

30 First, a PCR is conducted using the above-prepared primers (e.g., SEQ ID Nos.7 and 8) taking chromosomal DNA of the hyperthermophilic archaeon strain KOD1 as a template to amplify the DNA fragment. The DNA sequence (e.g., SEQ ID No.9) of the amplified fragment is determined and, after confirming that the originally set amino acid sequence is coded for, a Southern hybridization is conducted to the cleaved product of the chromosomal DNA with a restriction enzyme using said fragment as a probe. It is preferred that the approximate size of the fragment containing the aimed DNA polymerase gene is limited to about 4-7 Kbp.

35 Then DNA fragment of about 4-7 Kbp is recovered from the gel, a DNA library is prepared by *Escherichia coli* using said fragment and a colony hybridization is carried out using the above-mentioned PCR-amplified DNA fragment (e.g., SEQ ID No.9) to collect a clone strain.

The DNA polymerase gene of the strain KOD1 cloned in the present invention is composed of 5010 bases (estimated numbers of amino acids: 1670) (SEQ ID No.5).

40 Upon comparison with other DNA polymerases, there is a conserved region of αDNA polymerase which is an eukaryote type (Regions 1-5) in the gene of the present invention. In addition, there are EXO 1,2,3 which are 3'→5' exonuclease motives at the N terminal of said gene. In the conserved regions (Regions 1, 2) of the thermostable DNA polymerase gene derived from the hyperthermophilic archaeon strain KOD1, each of the intervening sequences is present and they are connected in a form where the open reading frame (ORF) is conserved.

45 When the thermostable DNA polymerase gene of the hyperthermophilic archaeon strain KOD1 is compared with Pfu DNA polymerase gene derived from *Pyrococcus furiosus* (Japanese Laid-Open Patent Publication Hei-05/328969) and with Tli (Vent) DNA polymerase gene derived from *Thermococcus litoralis* (Japanese Laid-Open Patent Publication He-06/7160) which are known enzymes, intervening sequence is present in the gene of the strain KOD1 of the present invention while there is no intervening sequence in the gene of the above-mentioned Pfu DNA polymerase and, in the 50 Tli DNA polymerase gene, there are two kinds of intervening sequences but they are present within Regions 2 and 3 which are conserved regions and that greatly differs from the location where the intervening sequence in the thermostable DNA polymerase gene of KOD1 strain of the present invention exists (see Fig. 7).

The gene of the present invention is a DNA which codes for the DNA polymerase derived from the hyperthermophilic archaeon strain KOD1. An example of said DNA contains a base sequence which codes for the amino acid sequence mentioned in SEQ ID No. 1 or 5. Further, such a DNA contains a base sequence mentioned in SEQ ID No. 5 or 6 or a part thereof.

In order to express the thermos table DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 of the present invention in *Escherichia coli*, the intervening sequences of 1374-2453 bp and 2708-4316 bp in the base sequence shown by SEQ ID No.5 are removed by means of a PCR gene fusion to construct a DNA polymerase gene

of a complete form. To be specific, a PCR is conducted on a cloned gene containing the intervening sequence by a combination of three pairs of primers to amplify the three fragments which are divided by the intervening sequence. In designing the primers used here, a part of the fragment which is to be bonded to its terminal is contained in its 5'-end. Then a PCR is conducted using the fragments to be bonded utilizing the duplicated sequence of the terminal whereby each of the fragments is bonded. Further PCR is conducted by the same manner using the resulting two kinds of fragments to give a DNA polymerase gene in a complete form containing no DNA polymerase gene derived from the strain KOD1 containing no intervening sequence.

Any vector may be used in the present invention so far as it makes cloning and expression of the thermostable DNA polymerase derived from KOD1 possible, e.g. a phage or plasmid. An example of the plasmid is a plasmid vector wherein an expression induced by T7 promoter is possible such as pET-8c. Other examples of the plasmid are pUC19, pBR322, pBluescript, pSP73, pGW7, pET3A and pET11C and so on. Examples of the phage are lambda gt11, lambda DASH and lambda ZapII and so on.

Examples of the host cell used in the present invention are Escherichia coli and yeasts. Examples of Escherichia coli are JM109, 101, XL1, PR1 and BL21(DE3)pysS and so on.

In the present invention, the gene coding for the thermostable DNA polymerase derived from the above-mentioned KOD1 is inserted into the above-mentioned vector to give a recombinant vector and the host cell is subjected to a transformation using said recombinant vector.

In the producing method of the present invention, the above-mentioned recombinant host cell is cultured whereby the thermostable DNA polymerase gene derived from the strain KOD1 is induced and expressed. The culture medium used for the culture of the recombinant host cell and the conditions therefor follow the conventional methods.

In a specific example, Escherichia coli which is transformed by pET-8c plasmid containing a DNA polymerase gene in a complete form containing no intervening sequence derived from the strain KOD1 is cultured, for example, in a TB medium whereby an induction treatment is conducted. It is preferred that the induction treatment of T7 promoter is carried out by addition of isopropylthio- β -D-galactoside.

The purifying method of the present invention includes, after culturing the recombinant host cells, a step wherein (a) recombinant host cells are collected, disintegrated and the cell extract is prepared and a step wherein (b) impure protein derived from the host cells is removed.

The thermostable DNA polymerase which is produced from the recombinant host cells is separated and recovered from the culture liquid by means of centrifugation or the like after culturing the host bacterial cells in a medium followed by inducing. After said bacterial cells are resuspended in a buffer, they are disintegrated by means of ultrasonic treatment, Dyno mill, French press, etc. Then a thermal treatment is conducted and the heat stable DNA polymerase is recovered from the supernatant fluid. In disintegrating the bacterial cells, ultrasonic treatment, Dyno mill and French press method are preferred.

A thermal treatment is preferred as one of the steps for removing the impure protein derived from the host cells. The condition for the thermal treatment is at 70°C or higher or, preferably, at 90° or higher. Other means for removing the impure protein are various chromatographic techniques.

Molecular weight of the thermostable DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 obtained as such is about 90 KDa (cf. Fig. 5).

When a polymerase chain reaction is conducted using said thermostable DNA polymerase, a sufficient amplification of the aimed DNA fragments is confirmed (cf. Fig. 6).

Now the present invention will be illustrated by referring partly to the drawings wherein:

Fig. 1 is a photographic picture of electrophoresis as a substitute for a drawing and shows the result of the measurement of the DNA extension rate of the KOD polymerase;

Fig. 2 is a photographic picture of electrophoresis as a substitute for a drawing and shows the comparison of the DNA extension rate of various thermostable DNA polymerases in which Fig. 2a shows the cases of KOD polymerase and Pfu polymerase while Fig. 2b shows the cases of Deep Vent polymerase and Taq polymerase;

Fig. 3 is a photographic picture of electrophoresis as a substitute for a drawing and shows the comparison of the PCR due to the difference in the reaction time of various thermostable DNA polymerase;

Fig. 4 shows the constructive charts of the recombinant expression vector;

Fig. 5 is a photographic picture of electrophoresis as a substitute for a drawing and shows the result of the measurement of molecular weight of the thermostable DNA polymerase derived from KOD1;

Fig. 6 is a photographic picture of electrophoresis as a substitute for a drawing and shows the result of the PCR by the thermostable DNA polymerase derived from KOD1; and

Fig. 7 shows drawings which show a comparison of the DNA polymerase gene derived from the hyperthermophilic archaeon strain KOD1 with the thermostable DNA polymerase gene derived from Pyrococcus furiosus and that derived from Thermococcus litoralis which are thought to be similar bacteria.

Example 1.

Cloning of DNA Polymerase Gene Derived from hyperthermophilic archaeon strain KOD1

5 The hyperthermophilic archaeon strain KOD1 isolated in Kodakara Island, Kagoshima was cultured at 95°C and then the bacterial cells were recovered. Chromosomal DNA of the hyperthermophilic archaeon strain KOD1 was prepared by a conventional method from the resulting bacterial cells.

10 Two kinds of primers (5'-GGATTAGTATAGTCCAATGGAAGGCGAC-3' [SEQ ID No.7] and 5'-GAGGGCGAAGTT-TATTCGAGCTT-3' [SEQ ID No.8]) were synthesized based upon the amino acid sequence at the conserved region of the DNA polymerase (Pfu polymerase) derived from Pyrococcus furiosus. A PCR was carried out using those two primers where the prepared chromosomal DNA was used as a template.

15 After the base sequence (SEQ ID No.9) of the PCR-amplified DNA fragment was determined and the amino acid sequence (SEQ ID No.10) was determined, a Southern hybridization was conducted using said amplified DNA fragment to the product of the strain KOD1 chromosomal DNA treated with a restriction enzyme whereby the size of the fragment coding for the DNA polymerase was calculated (about 4-7 Kbp). Further, the DNA fragment of this size was recovered from agarose gel, inserted into a plasmid pBS (manufactured by Stratagene) and Escherichia coli (E. coli JM 109) was transformed by this mixture to prepare a library.

20 A colony hybridization was conducted using a probe (SEQ ID No.9) used for the Southern hybridization to obtain a clone strain (E. coli JM109/pBSK0D1) which is thought to contain the DNA polymerase gene derived from strain KOD1.

Example 2.

Determination of Base Sequence of the cloned Fragment

25 A plasmid pBSKOD1 was recovered from the clone strain E. coli JM109/pBSKOD1 obtained in Example 1 and its base sequence (SEQ ID No.5) was determined by a conventional method. Further, the amino acid sequence was derived from the determined base sequence. The DNA polymerase gene derived from KOD1 strain comprised 5010 bases wherein 1670 amino acids were coded.

Example 3.

Construction of Recombinant Expression Vector

35 In order to prepare a complete polymerase gene, the intervening sequence parts at two places (1374-2453 bp and 2708-4316 bp) were removed by a PCR fusion method. In the PCR fusion method, three pairs of primers (SEQ ID Nos.11-16) were combined using a primer recovered from the clone strain as a template and a PCR was conducted for each of them to amplify three fragments wherefrom the intervening sequences were removed. At that time, the primer used for the PCR was designed in such a manner that the side which binds to another fragment has the same sequence as the binding partner has. In addition, a design was conducted in such a manner that different restriction enzyme sites (EcoRV at the N-terminus while BamHI at the C-terminus) were created at both ends.

40 After that, among the PCR-amplified fragments, that which is located at the central part of the structure and that which is located at the N-terminal side are mixed and a PCR was conducted using each of the fragments as a primer. At the same time, the fragment located at the central part of the structure and that located at the C-terminal side are mixed and a PCR was conducted using each of the fragments as a primer. Two kinds of fragments obtained as such were subjected to a PCR once again to give gene fragments in a complete form having no intervening sequence, having EcoRV and BamHI sites at the N- and C-termini, respectively and coding for the DNA polymerase derived from strain KOD1.

45 Further, said gene was subcloned using an expression vector which can be induced by T7 promoter, an NcoI/BamHI site of pET-8c and the previously-created restriction enzyme site to give a recombinant expression vector (pET-pol).

Example 4.

Expression and Purification of DNA Polymerase Derived from KOD1

55 Escherichia coli (BL21(DE3)) was transformed using a recombinant expression vector (pET-pol) obtained in Example 3, the resulting transformant was cultured in a TB medium (mentioned in Molecular Cloning, p.A.2, 1989) and, one hour before collecting the bacterial cells, an induction treatment of T7 promoter was conducted by addition of isopropylthio- β -D-galactopyranoside. Bacterial cells were recovered from the cultured liquid by means of centrifugation. They

were resuspended in a buffer and disintegrated by an ultrasonic treatment to give a cell extract. In order to remove the impure protein derived from the host cells, the disintegrated cell solution was treated at 94°C for 20 minutes whereby the impure protein derived from the host cells was isolated by centrifugation to give a thermostable DNA polymerase derived from strain KOD1.

5 The *Eschericia coli* BL21 (DE3) pER-pol was deposited on April 22, 1996 under the Budapest Treaty at National Institute of Bioscience and Human-Technology Agency of Industrial Science and Technology (1-3, Higashi 1 chome Tsukuba-shi Ibaraki-ken 305, JAPAN) in accordance with the Budapest Treaty under the accession number FERM BP-5513.

10 Example 5.

Purification of Thermostable DNA Polymerase Derived from KOD1

15 Molecular weight of the thermostable DNA polymerase derived from KOD1 obtained in Example 4 was calculated by means of an SDS-PAGE method whereby it was found to be about 86-92 kDa (Fig. 5). Further, a PCR was conducted using the thermostable DNA polymerase derived from KOD1 obtained in Example 4 and the known template primer whereupon a DNA fragment which was to be a target was confirmed (Fig. 6) by the same manner as in the case where the thermostable DNA polymerase derived from *Thermococcus litoralis* was used and a high thermostable DNA polymerase activity was confirmed.

20 Comparative Example 1.

Comparison with the Thermostable DNA Polymerase Gene Derived from *Pyrococcus furiosus* or from *Thermococcus litoralis* which are to be Similar to the Hyperthermophilic archaeon strain KOD1 of the Present Invention.

25 Amino acid sequences were estimated from the DNA sequences of the DNA polymerase gene derived from the hyperthermophilic archaeon strain KOD1 of the present invention (SEQ ID No.6), the thermostable DNA polymerase gene derived from *Pyrococcus furiosus* (Japanese Laid-Open Patent Publication Hei-5/328969) and the thermostable DNA polymerase gene derived from *Thermococcus litoralis* (Japanese Laid-Open Patent Publication Hei-6/7160) and were compared and investigated.

30 In the DNA polymerase derived from KOD1 of the present invention, there were regions 1-5 which were the conserved regions of α DNA polymerase of an eukaryotic type. Further, there were EXO1, 2 and 3 which were 3'- \rightarrow 5' exonuclease motives at the N-terminal side. However, in each of the Region 1 and Region 2 which were the α DNA polymerase conserved regions, there were intervening sequences IVS-A and IVS-B (refer to Fig. 7).

35 On the other hand, in Pfu polymerase which is a thermostable DNA polymerase derived from *Pyrococcus furiosus*, there was no intervening sequence. In the case of Vent polymerase which is a thermostable DNA polymerase derived from *Thermococcus litoralis*, there were the intervening sequences (IVS1 and IVS2) in the α DNA polymerase conserved regions (Region 2 and Region 3) (see Fig.7).

40 Example 6.

Measurement of DNA Extension Rate of the DNA Polymerase Derived from Hyperthermophilic archaeon strain KOD1

45 DNA prepared by annealing the M13mp18DNA with M13P7 primer having a base sequence as mentioned in SEQ ID No.2 was used as a substrate and the rate of synthesizing the DNA in a reaction buffer solution [20 mM Tris-HCl (pH 7.5 at 75°C), 10 mM KCl, 6 mM (NH₄)₂SO₄, 2 mM MgCl₂, 0.1% Triton X-100 and 10 μ g/ml nuclease-free BSA] containing the DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 manufactured in Examples 1-5 was investigated for the reaction time of 20, 40, 60, 80 and 100 seconds (Fig. 1) or 40, 60, 80 and 100 seconds (Fig. 2). The results are given in Fig. 1 and in Fig. 2.

50 A part of the DNA sample during the elongation reaction was taken out for each reaction time and was added to a reaction stopping solution (60 mM EDTA, 60 μ M NaOH, 0.1% BPB and 30% glycerol) in the same amount.

The DNA samples obtained in the above process were separated and analyzed by means of an alkaline agarose electrophoresis and the size of the synthesized DNA was checked.

55 1, 2, 3, 4 and 5 in Fig. 1 show the results of the reactions for 0.3 minute (20 seconds), 0.7 minute (40 seconds), 1 minute (60 seconds), 1.3 minutes (80 seconds) and 1.7 minutes (100 seconds), respectively. It is apparent from Fig. 1 that the DNA extension rate of the DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 was 105 bases/second.

1, 2, 3 and 4 in Fig. 2 show the results of the reaction for 0.7 minute (40 seconds), 1 minute (60 seconds), 1.3 minutes (80 seconds) and 1.7 minutes (100 seconds), respectively. It is apparent from Fig. 2 that the DNA extension rate

of the DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 was 138 bases/second.

On the other hand, the DNA synthesizing rate of each of Pfu polymerase (Stratgene), Deep Vent polymerase (New England Biolab) and Taq polymerase (Takara Shuzo) was measured by the same manner in each of the buffers therefor (Fig. 2a and Fig. 2b). The DNA extension rates of those DNA polymerases were 24.8 bases/second for Pfu polymerase, 23.2 bases/second for Deep Vent polymerase and 61.0 bases/second for Taq polymerase.

From the above results, it was suggested that the DNA extension rate of the DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 was about six-fold of those of Pfu polymerase and Deep Vent polymerase and about two-fold of that of Taq polymerase.

10 Example 7.

Measurement of Fidelity of the DNA Polymerase Derived from the Hyperthermophilic archaeon strain KOD1 in the Reaction for the Synthesis of DNA

15 The error rate in the DNA synthesis was measured by the method of Kunkel (Kunkel, 1985, Journal of Biological Chemistry, 260, 5787-5796). In this method, a DNA synthesis reaction was conducted using a DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 manufactured in Examples 1-5 using an M13mp18DNA having a gap at a lacZ part containing a part of the genes coding for β -galactosidase as a substrate and transfected to *E. coli* JM109 in an NZY medium containing 5-bromo-4-chloro-3-indolyl- β -D-galactoside and isopropyl-thio- β -D-galactoside using an

20 M13mp18DNA in which lacZ part was double-stranded.

When β -galactosidase wherein a function is lost or lowered was expressed due to a reading error or a frame shift during the synthetic reaction of DNA, it is not possible to utilize 5-bromo-4-chloro-3-indolyl- β -D-galactoside whereupon the color of plaque becomes colorless or light blue. On the other hand, when there is no error in the synthesized DNA and a complete β -galactosidase was expressed, plaque becomes blue. The rate of induction of error was measured in

25 the DNA synthesis from the rate of the sum of colorless and light blue plaque to the total plaque.

The rate of induction of error in the DNA synthesis was also measured for Pfu polymerase (Stratgene), Taq polymerase (Takara Shuzo) and delta Tth polymerase (Toyobo) which were made to react by the same manner.

Further, the rate of induction of error in the DNA synthesis was also measured for a mixture of Taq polymerase and Pfu polymerase. The results are given in Table 1.

30

Table 1

Measurement of Fidelity in the Reaction of DNA Synthesis of DNA Polymerase Derived from Hyperthermophilic archaeon strain KOD1					
Enzyme	Light Blue	White	Mutant	Total	Mutant Frequency(10^{-4})
KOD1 pol.	12	11	23	6619	37.7
Pfu	15	15	30	7691	39.0
Taq	30	24	54	4141	130
▲Tth	70	45	115	7375	156
Taq/Pfu(20:1)	10	20	30	4238	63.7
Taq/Pfu(50:1)	10	13	23	4489	53.5

35

40

45

It is apparent from Table 1 that the fidelity of the DNA polymerase derived from hyperthermophilic archaeon strain KOD1 in the DNA synthesis reaction is suggested to be superior to Taq polymerase and same as Pfu polymerase. In addition, a mixture of Taq polymerase and Pfu polymerase exhibits a medium fidelity that it is superior to Taq polymerase and inferior to Pfu polymerase.

Example 8.

55 Comparison in PCR of Various Thermostable DNA Polymerases by the Difference in the Reaction Time

lambda-DNA (3 μ g) was used as a target nucleic acid; oligo-nucleotides having a sequence as mentioned in SEQ ID Nos. 3 and 4 were used as primers; and a buffer containing 20 mM Tris-HCl (pH 7.5 at 75°C), 10 mM KCl, 6 mM $(\text{NH}_4)_2\text{SO}_4$, 2 mM MgCl₂, 0.1% Triton X-100, 10 μ g/ml BSA and 200 μ M dNTPs was used as a buffer. DNA polymerase

EP 0 745 675 A2

derived from hyperthermophilic archaeon strain KOD1 (KOD polymerase), Taq polymerase which is widely used for PCR and Pfu polymerase which exhibits 3'-5' exonuclease activity were also used as the thermostable DNA polymerases. The used titer of each polymerase was 2 units.

5 A PCR amplification reaction was conducted using a DNA Thermal Cycler (Perkin-Elmer) in a schedule wherein a cycle comprising 94°C/20 seconds and 68°C/x second (x: reaction time) was repeated for 30 times. In the case of the DNA polymerase derived from the hyperthermophilic archaeon strain KOD1 (KOD polymerase), amplification of the target DNA was confirmed by conducting 30 cycles of 94°C/20 seconds-68°C/1 second while, in the case of Taq polymerase, amplification of DNA was first confirmed by conducting 30 cycles of 94°C/20 seconds-68°C/10 seconds. In the case of Pfu polymerase, amplification of DNA was at least confirmed by conducting 30 cycles of 94°C/20 seconds-68°C/1 minute. The results are given in Fig. 3.

10 In the present invention, it is possible to amplify the DNA with a high fidelity within a short reaction time when a DNA polymerase derived from hyperthermophilic archaeon strain KOD1 which is a thermostable DNA polymerase having at least 30 bases/second of DNA extension rate and having a 3'-5' exonuclease activity. When this method is carried out, it is possible to improve the simplicity and convenience. In addition, when such kind of thermostable DNA polymerase 15 having both high extension rate (at least 30 bases/second) which has not been available yet and 3'-5' exonuclease activity is used, it is possible to shorten the time for the primer extention reaction and to amplify the relatively big product with a high fidelity.

20

25

30

35

40

45

50

55

SEQUENCE LISTING

5 SEQ ID No.1

Length: 774 base pairs

10 Type: amino acid

Topology: linear

15 Type: protein

Sequence Description:

20 Met Ile Leu Asp Thr Asp Tyr Ile Thr Glu Asp Gly Lys Pro Val Ile

1

5

10

15

25 Arg Ile Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg

20

25

30

30 Thr Phe Glu Pro Tyr Phe Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile

35

40

45

35 Glu Glu Val Lys Lys Ile Thr Ala Glu Arg His Gly Thr Val Val Thr

50

55

60

40 Val Lys Arg Val Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Val

65

70

75

80

45 Glu Val Trp Lys Leu Tyr Phe Thr His Pro Gln Asp Val Pro Ala Ile

85

90

95

50 Arg Asp Lys Ile Arg Glu His Gly Ala Val Ile Asp Ile Tyr Glu Tyr

100

105

110

55 Asp Ile Pro Phe Ala Lys Arg Tyr Leu Ile Asp Lys Gly Leu Val Pro

115

120

125

55 Met Glu Gly Asp Glu Glu Leu Lys Met Leu Ala Phe Asp Ile Gln Thr

Pro Asn Lys Pro Asp Glu Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr
 5 370 375 380
 Glu Gly Gly Tyr Val Lys Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile
 385 390 395 400
 10 Val Tyr Leu Asp Phe Arg Ser Leu Tyr Pro Ser Ile Ile Ile Thr His
 405 410 415
 Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu Tyr Asp
 15 420 425 430
 Val Ala Pro Gln Val Gly His Arg Phe Cys Lys Asp Phe Pro Gly Phe
 435 440 445
 20 Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys Ile Lys
 450 455 460
 Lys Lys Met Lys Ala Thr Ile Asp Pro Ile Glu Arg Lys Leu Leu Asp
 25 465 470 475 480
 Tyr Arg Gln Arg Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr
 485 490 495
 30 Tyr Gly Tyr Ala Arg Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser
 500 505 510
 35 Val Thr Ala Trp Gly Arg Glu Tyr Ile Thr Met Thr Ile Lys Glu Ile
 515 520 525
 Glu Glu Lys Tyr Gly Phe Lys Val Ile Tyr Ser Asp Thr Asp Gly Phe
 40 530 535 540
 Phe Ala Thr Ile Pro Gly Ala Asp Ala Glu Thr Val Lys Lys Ala
 545 550 555 560
 45 Met Glu Phe Leu Asn Tyr Ile Asn Ala Lys Leu Pro Gly Ala Leu Glu
 565 570 575
 Leu Glu Tyr Glu Gly Phe Tyr Lys Arg Gly Phe Phe Val Thr Lys Lys
 50 580 585 590
 Lys Tyr Ala Val Ile Asp Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu

EP 0 745 675 A2

130 135 140
5 Leu Tyr His Glu Gly Glu Glu Phe Ala Glu Gly Pro Ile Leu Met Ile
145 150 155 160
10 Ser Tyr Ala Asp Glu Gly Ala Arg Val Ile Thr Trp Lys Asn Val
165 170 175
15 Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Arg Glu Met Ile Lys
180 185 190
19 Arg Phe Leu Arg Val Val Lys Glu Lys Asp Pro Asp Val Leu Ile Thr
195 200 205
20 Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys Lys Arg Cys Glu
210 215 220
25 Lys Leu Gly Ile Asn Phe Ala Leu Gly Arg Asp Gly Ser Glu Pro Lys
225 230 235 240
25 Ile Gln Arg Met Gly Asp Arg Phe Ala Val Glu Val Lys Gly Arg Ile
245 250 255
30 His Phe Asp Leu Tyr Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr
260 265 270
35 Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Phe Gly Gln Pro Lys Glu
275 280 285
35 Lys Val Tyr Ala Glu Glu Ile Thr Pro Ala Trp Glu Thr Gly Glu Asn
290 295 300
40 Leu Glu Arg Val Ala Arg Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr
305 310 315 320
45 Glu Leu Gly Lys Glu Phe Leu Pro Met Glu Ala Gln Leu Ser Arg Leu
325 330 335
50 Ile Gly Gln Ser Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu
340 345 350
50 Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala
355 360 365

EP 0 745 675 A2

595 600 605
Glu Ile Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala
5 610 615 620
Arg Val Leu Glu Ala Leu Leu Lys Asp Gly Asp Val Glu Lys Ala Val
10 625 630 635 640
Arg Ile Val Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro
645 650 655
15 Pro Glu Lys Leu Val Ile His Glu Gln Ile Thr Arg Asp Leu Lys Asp
660 665 670
Tyr Lys Ala Thr Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala
20 675 680 685
Arg Gly Val Lys Ile Arg Pro Gly Thr Val Ile Ser Tyr Ile Val Leu
690 695 700
25 Lys Gly Ser Gly Arg Ile Gly Asp Arg Ala Ile Pro Phe Asp Glu Phe
705 710 715 720
Asp Pro Thr Lys His Lys Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln
30 725 730 735
Val Leu Pro Ala Val Glu Arg Ile Leu Arg Ala Phe Gly Tyr Arg Lys
740 745 750
35 Glu Asp Leu Arg Tyr Gln Lys Thr Arg Gln Val Gly Leu Ser Ala Trp
755 760 765
40 Leu Lys Pro Lys Gly Thr
770

45

50

55

50 SEQ ID No.2

Length: 24 base pairs

5 Type: nucleic acid

Topology: linear

10 Type: synthetic DNA

Sequence Description:

CGCCAGGGTT TTCCAGTCA CGAC

24

55 SEQ ID No.3

Length: 20 base pairs

20 Type: nucleic acid

Topology: linear

Molecular Type: synthetic DNA

25 Sequence Description:

GGGCAGGCGAC CTCGCAGGTT

20

SEQ ID No.4

30 Length: 24 base pairs

Type: nucleic acid

Topology: linear

35 Molecular Type: synthetic DNA

Sequence Description:

GCCCATAATA ATCTGCCGGT CAAT

24

40 SEQ ID No.5

Length: 5342 base pairs

45 Type: nucleic acid (DNA)

Strandedness: double

Topology: linear

50 Molecular Type: cDNA

Source: hyperthermophilic archaeon

55

EP 0 745 675 A2

5 Strain name: KOD1
 Characteristics: { 156-5165 P CDS
 { 1374-2453 intervening sequence
 { 2708-4316 intervening sequence

Sequence Description:

10 GCTTGAGGGC CTGCGTTAT CGGACGTTGC AGTMTGCCGC TACTCAAAGA TGCCCGTTTT 60
 ATAACGGAGA AAAATGGGA GCTATTACGA TCTCTCCTTC ATGTGGGTT TACAATAAAC 120
 CCTGGATTGT TCTACAAGAT TATGGGGAT GAAAG ATG ATC CTC GAC ACT GAC 173
 Met Ile Leu Asp Thr Asp
 15 1 5
 TAC ATA ACC GAG GAT GGA AAG CCT GTC ATA AGA ATT TTC AAG AAG GAA 221
 Tyr Ile Thr Glu Asp Gly Lys Pro Val Ile Arg Ile Phe Lys Lys Glu
 20 10 15 20
 AAC CCC GAG TTT AAG ATT GAG TAC GAC CGG ACT TTT GAA CCC TAC TTC 269
 Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg Thr Phe Glu Pro Tyr Phe
 25 25 30 35
 TAC GCC CTC CTG AAG GAC GAT TCT GCC ATT GAG GAA GTC AAG AAG ATA 317
 Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile Glu Glu Val Lys Lys Ile
 30 40 45 50
 ACC GCC GAG AGG CAC GGG ACG GTT GTA ACG GTT AAG CGG GTT GAA AAG 365
 Thr Ala Glu Arg His Gly Thr Val Val Thr Val Lys Arg Val Glu Lys
 35 55 60 65 70
 GTT CAG AAG AAG TTC CTC GGG AGA CCA GTT GAG GTC TGG AAA CTC TAC 413
 Val Gln Lys Lys Phe Leu Gly Arg Pro Val Glu Val Trp Lys Leu Tyr
 40 75 80 85
 TTT ACT CAT CCG CAG GAC GTC CCA CGG ATA AGG GAC AAG ATA CGA CAG 461
 Phe Thr His Pro Gln Asp Val Pro Ala Ile Arg Asp Lys Ile Arg Glu
 45 90 95 100
 CAT GGA GCA GTT ATT GAC ATC TAC GAC TAC GAC ATA CCC TTC GCC AAG 509
 His Gly Ala Val Ile Asp Ile Tyr Glu Tyr Asp Ile Pro Phe Ala Lys
 50 105 110 115

EP 0 745 675 A2

5	CCC TAC CTC ATA GAC AAG GGA TTA GTG CCA ATG GAA GGC GAC GAG GAG	557
	Arg Tyr Leu Ile Asp Lys Gly Leu Val Pro Met Glu Gly Asp Glu Glu	
	120 125 130	
10	CTG AAA ATG CTC GCC TTC GAC ATT CAA ACT CTC TAC CAT GAG GGC GAG	605
	Leu Lys Met Leu Ala Phe Asp Ile Gln Thr Leu Tyr His Glu Gly Glu	
	135 140 145 150	
15	GAG TTC GCC GAG CGG CCA ATC CTT ATG ATA ACC TAC GCC GAC GAG GAA	653
	Glu Phe Ala Glu Gly Pro Ile Leu Met Ile Ser Tyr Ala Asp Glu Glu	
	155 160 165	
20	GGG GCC AGG GTG ATA ACT TGG AAG AAC GTG GAT CTC CCC TAC GTT GAC	701
	Gly Ala Arg Val Ile Thr Trp Lys Asn Val Asp Leu Pro Tyr Val Asp	
	170 175 180	
25	GTC GTC TCG ACG GAG AGG GAG ATG ATA AAG CGC TTC CTC CGT GTT GTG	749
	Val Val Ser Thr Glu Arg Glu Met Ile Lys Arg Phe Leu Arg Val Val	
	185 190 195	
30	AAG GAG AAA GAC CCG GAC GTT CTC ATA ACC TAC AAC GGC GAC AAC TTC	797
	Lys Glu Lys Asp Pro Asp Val Leu Ile Thr Tyr Asn Gly Asp Asn Phe	
	200 205 210	
35	GAC TTC CCC TAT CTG AAA AAG CGC TGT GAA AAG CTC CGA ATA AAC TTC	845
	Asp Phe Ala Tyr Leu Lys Lys Arg Cys Glu Lys Leu Gly Ile Asn Phe	
	215 220 225 230	
40	GCC CTC GGA AGG GAT GGA AGC GAG CCG AAG ATT CAG AGG ATG GGC GAC	893
	Ala Leu Gly Arg Asp Gly Ser Glu Pro Lys Ile Gln Arg Met Gly Asp	
	235 240 245	
45	AGG TTT GCC GTC GAA GTG AAG GGA CGG ATA CAC TTC GAT CTC TAT CCT	941
	Arg Phe Ala Val Glu Val Lys Gly Arg Ile His Phe Asp Leu Tyr Pro	
	250 255 260	
50	GTG ATA AGA CGG ACG ATA AAC CTG CCC ACA TAC ACG CTT GAG GCC GTT	989
	Val Ile Arg Arg Thr Ile Asn Leu Pro Thr Tyr Thr Leu Glu Ala Val	

55

EP 0 745 675 A2

	265	270	275	
5	TAT CAA GCC GTC TTC GGT CAG CCG AAG GAG AAG GTT TAC GCT GAG GAA			1037
	Tyr Glu Ala Val Phe Gly Gln Pro Lys Glu Lys Val Tyr Ala Glu Glu			
	280	285	290	
10	ATA ACA CCA GCC TGG GAA ACC GGC GAG AAC CTT GAG AGA GTC GCC CGC			1085
	Ile Thr Pro Ala Trp Glu Thr Gly Glu Asn Leu Glu Arg Val Ala Arg			
	295	300	305	310
15	TAC TCG ATG GAA GAT GCG AAG GTC ACA TAC GAG CTT GGG AAG GAG TTC			1133
	Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr Glu Leu Gly Lys Glu Phe			
	315	320	325	
20	CTT CCG ATG GAG CCC CAG CTT TCT CGC TTA ATC GGC CAG TCC CTC TGG			1181
	Leu Pro Met Glu Ala Gln Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp			
	330	335	340	
25	GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC			1229
	Asp Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu			
	345	350	355	
30	AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA			1277
	Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu			
	360	365	370	
35	AAG GAG CTG GCC AGA AGA CGG CAG ACC TAT GAA GGA GGC TAT GTC AAA			1325
	Lys Glu Leu Ala Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys			
	375	380	385	390
40	GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA			1373
	Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg			
	395	400	405	
45	TGC CAT CCA GCC GAT ACG AAC GTT GTC GTC AAC GGG AAG GGG ATT ATA			1421
	Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile			
	410	415	420	
50	AAC ATC ACC GAG GTT CAG GAA GGT GAC TAT GTC CTT GGG ATT GAC GGC			1465

5	Asn Ile Ser Glu Val Gin Glu Gly Asp Tyr Val Leu Gly Ile Asp Gly	425	430	435	
	TGG CAG AGA GTT AGA AAA GTA TGG GAA TAC GAC TAC AAA GGG GAG CTT				1517
10	Trp Gin Arg Val Arg Lys Val Trp Glu Tyr Asp Tyr Lys Gly Glu Leu	440	445	450	
	GTA AAC ATA AAC CGG TTA AAG TGT ACG CCC AAT CAT AAG CTT CCC GTT				1565
15	Val Asn Ile Asn Gly Leu Lys Cys Thr Pro Asn His Lys Leu Pro Val	455	460	465	470
	GTT ACA AAG AAC GAA CGA CAA ACG AGA ATA AGA GAC AGT CTT GCT AAG				1613
20	Val Thr Lys Asn Glu Arg Gin Thr Arg Ile Arg Asp Ser Leu Ala Lys	475	480	485	
	TCT TTC CTT ACT AAA AAA GTT AAG GGC AAG ATA ATA ACC ACT CCC CTT				1661
25	Ser Phe Leu Thr Lys Lys Val Lys Gly Lys Ile Ile Thr Thr Pro Leu	490	495	500	
	TTC TAT GAA ATA GGC AGA GCG ACA AGT GAG AAT ATT CCA GAA GAA GAG				1709
30	Phe Tyr Glu Ile Gly Arg Ala Thr Ser Glu Asn Ile Pro Glu Glu Glu	505	510	515	
	GTT CTC AAG GGA GAG CTC GCT GGC ATA CTA TTG GCT GAA GGA ACG CTC				1757
35	Val Leu Lys Gly Glu Leu Ala Gly Ile Leu Leu Ala Glu Gly Thr Leu	520	525	530	
	TTG AGG AAA GAC GTT GAA TAC TTT GAT TCA TCC CGC AAA AAA CGG AGG				1805
40	Leu Arg Lys Asp Val Glu Tyr Phe Asp Ser Ser Arg Lys Lys Arg Arg	535	540	545	550
	ATT TCA CAC CAG TAT CGT GTT GAG ATA ACC ATT GGG AAA GAC GAG GAG				1853
45	Ile Ser His Gin Tyr Arg Val Glu Ile Thr Ile Gly Lys Asp Glu Glu	555	560	565	
	GAG TTT AGG GAT CGT ATC ACA TAC ATT TTT GAG CGT TTG TTT GGG ATT				1901
50	Glu Phe Arg Asp Arg Ile Thr Tyr Ile Phe Glu Arg Leu Phe Gly Ile	570	575	580	

	ACT CCA AGC ATC TCG GAG AAG AAA GGA ACT AAC GCA GTC ACA CTC AAA		1949
5	Thr Pro Ser Ile Ser Glu Lys Lys Gly Thr Asn Ala Val Thr Leu Lys		
	585	590	595
	GTT GCG AAG AAG AAT GTT TAT CTT AAA GTC AAG GAA ATT ATG GAC AAC		1997
10	Val Ala Lys Lys Asn Val Tyr Leu Lys Val Lys Glu Ile Met Asp Asn		
	600	605	610
	ATA GAG TCC CTA CAT GCC CCC TCG GTT CTC ACG GGA TTC TTC GAA GGC		2045
15	Ile Glu Ser Leu His Ala Pro Ser Val Leu Arg Gly Phe Phe Glu Gly		
	615	620	625
	GAC GCT TCA GTA AAC AGG GTT AGG AGG ACT ATT GTT GCA ACC CAG GGT		2093
20	Asp Gly Ser Val Asn Arg Val Arg Arg Ser Ile Val Ala Thr Gln Gly		
	635	640	645
	ACA AAG AAC GAG TGG AAG ATT AAA CTG GTG TCA AAA CTG CTC TCC CAG		2141
25	Thr Lys Asn Glu Trp Lys Ile Lys Leu Val Ser Lys Leu Leu Ser Gln		
	650	655	660
	CTT GGT ATC CCT CAT CAA ACG TAC ACG TAT CAG TAT CAG GAA AAT CGG		2189
30	Leu Gly Ile Pro His Gln Thr Tyr Thr Tyr Gln Tyr Gln Glu Asn Gly		
	665	670	675
	AAA GAT CGG AGC AGG TAT ATA CTG GAG ATA ACT GGA AAG GAC CGA TTG		2237
35	Lys Asp Arg Ser Arg Tyr Ile Leu Glu Ile Thr Gly Lys Asp Gly Leu		
	680	685	690
	ATA CTG TTC CAA ACA CTC ATT GGA TTC ATC AGT GAA AGA AAG AAC GCT		2285
40	Ile Leu Phe Gln Thr Leu Ile Gly Phe Ile Ser Glu Arg Lys Asn Ala		
	695	700	705
	CTG CTT AAT AAC GCA ATA TCT CAG AGG GAA ATG AAC AAC TTG GAA AAC		2333
45	Leu Leu Asn Lys Ala Ile Ser Gln Arg Glu Met Asn Asn Leu Glu Asn		
	715	720	725
	AAT CGA TTT TAC AGG CTC AGT GAA TTC AAT GTC AGC ACG GAA TAC TAT		2381
50	Asn Gly Phe Tyr Arg Leu Ser Glu Phe Asn Val Ser Thr Glu Tyr Tyr		

EP 0 745 675 A2

	730	735	740	
5	GAG GGC AAG GTC TAT GAC TTA ACT CTT GAA GGA ACT CCC TAC TAC TTT			2429
	Glu Gly Lys Val Tyr Asp Leu Thr Leu Glu Gly Thr Pro Tyr Tyr Phe			
10	745	750	755	
	GCC AAT GGC ATA TTG ACC CAT AAC TCC CTG TAC CCC TCA ATC ATC ATC			2477
	Ala Asn Gly Ile Leu Thr His Asn Ser Leu Tyr Pro Ser Ile Ile Ile			
15	760	765	770	
	ACC CAC AAC GTC TCG CCG GAT ACG CTC AAC AGA GAA GGA TGC AAG GAA			2525
	Thr His Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu			
20	775	780	785	790
	TAT GAC GTT GCC CCA CAG GTC GGC CAC CGC TTC TGC AAG GAC TTC CCA			2573
	Tyr Asp Val Ala Pro Gln Val Gly His Arg Phe Cys Lys Asp Phe Pro			
25	795	800	805	
	GGA TTT ATC CCG AGC CTG CTT GGA GAC CTC CTA GAG GAG AGG CAG AAG			2621
	Gly Phe Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys			
30	810	815	820	
	ATA AAG AAG AAG ATG AAG GCC ACG ATT GAC CCG ATC GAG AGG AAG CTC			2669
	Ile Lys Lys Lys Met Lys Ala Thr Ile Asp Pro Ile Glu Arg Lys Leu			
35	825	830	835	
	CTC GAT TAC AGG CAG AGG GCC ATC AAG ATC CTG GCA AAC AGC ATC CTA			2717
	Leu Asp Tyr Arg Gln Arg Ala Ile Lys Ile Leu Ala Asn Ser Ile Leu			
40	840	845	850	
	CCC GAG GAA TGG CTT CCA GTC CTC GAG GAA GGG GAG GTT CAC TTC GTC			2765
	Pro Glu Glu Trp Leu Pro Val Leu Glu Glu Gly Glu Val His Phe Val			
45	855	860	865	870
	AGG ATT GGA GAG CTC ATA GAC CCG ATG ATG GAG GAA AAT GCT GGG AAA			2813
	Arg Ile Gly Glu Leu Ile Asp Arg Met Met Glu Glu Asn Ala Gly Lys			
50	875	880	885	
	CTA AAG AGA GAG CGC GAC ACC GAA CTG CTT GAG GTC AGT GGG CTT GAA			2861

55

	Val Lys Arg Glu Gly Glu Thr Glu Val Leu Glu Val Ser Gly Leu Glu			
5	890	895	900	
	GTC CCG TCC TTT AAC AGC AGA ACT AAC AAG GCC GAG CTC AAG AGA CTA			2909
	Val Pro Ser Phe Asn Arg Arg Thr Asn Lys Ala Glu Leu Lys Arg Val			
10	905	910	915	
	AAG GCC CTG ATT AGG CAC GAT TAT TCT CCC AAG GTC TAC ACC ATC AGA			2957
	Lys Ala Leu Ile Arg His Asp Tyr Ser Gly Lys Val Tyr Thr Ile Arg			
15	920	925	930	
	CTG AAG TCG GGG AGG AGA ATA AAG ATA ACC TCT GGC CAC AGC CTC TTC			3005
	Leu Lys Ser Gly Arg Arg Ile Lys Ile Thr Ser Gly His Ser Leu Phe			
20	935	940	945	950
	TCT CTG AGA AAC GGG GAG CTC GTT GAA GTT ACG GGC GAT GAA CTA AAG			3053
	Ser Val Arg Asn Gly Glu Leu Val Glu Val Thr Gly Asp Glu Leu Lys			
25	955	960	965	
	CCA GGT GAC CTC GTT GCA GTC CCG CCG AGA TTG GAG CTT CCT GAG AGA			3101
	Pro Gly Asp Leu Val Ala Val Pro Arg Arg Leu Glu Leu Pro Glu Arg			
30	970	975	980	
	AAC CAC CTG CTG AAC CTC GTT GAA CTG CTC CTT GGA ACG CCA GAA GAA			3149
	Asn His Val Leu Asn Leu Val Glu Leu Leu Leu Gly Thr Pro Glu Glu			
35	985	990	995	
	GAA ACT TTG GAC ATC GTC ATG ACG ATC CCA GTC AAG GGT AAG AAC AAC			3197
	Glu Thr Leu Asp Ile Val Met Thr Ile Pro Val Lys Gly Lys Lys Asn			
40	1000	1005	1010	
	TTC TTT AAA GGG ATG CTC AGG ACT TTG CGC TGG ATT TTC GGA GAG GAA			3245
	Phe Phe Lys Gly Met Leu Arg Thr Leu Arg Trp Ile Phe Gly Glu Glu			
45	1015	1020	1025	1030
	AAG AGG CCC AGA ACC GCG AGA CGC TAT CTC AGG CAC CTT GAG GAT CTG			3293
	Lys Arg Pro Arg Thr Ala Arg Arg Tyr Leu Arg His Leu Glu Asp Leu			
50	1035	1040	1045	

	1195	1200	1205	
5	GAG AAC AAG AGG ATT CCC GAG TTC GTC TTC ACG TCC CCG AAA GGG GTT			3821
	Glu Asn Lys Arg Ile Pro Glu Phe Val Phe Thr Ser Pro Lys Gly Val			
	1210	1215	1220	
10	CGG CTG GCC TTC CTT GAG GGG TAC TCA TCG GCG ATG GCG ACG TCC ACC			3869
	Arg Leu Ala Phe Leu Glu Gly Tyr Ser Ser Ala Met Ala Thr Ser Thr			
	1225	1230	1235	
15	GAA CAA GAG ACT CAG GCT CTC AAC GAA AAG CGA GCT TTA GCG AAC CAG			3917
	Glu Gln Glu Thr Gln Ala Leu Asn Glu Lys Arg Ala Leu Ala Asn Gln			
	1240	1245	1250	
20	CTC GTC CTC CTC TTG AAC TCG GTG GGG GTC TCT GCT GTC AAA CTT GGG			3965
	Leu Val Leu Leu Asn Ser Val Gly Val Ser Ala Val Lys Leu Gly			
	1255	1260	1265	1270
25	CAC GAC AGC GGC GTT TAC AGG GTC TAT ATA AAC GAG GAG CTC CCG TTC			4013
	His Asp Ser Gly Val Tyr Arg Val Tyr Ile Asn Glu Glu Leu Pro Phe			
	1275	1280	1285	
30	GTC AAG CTG GAC AAG AAA AAG AAC GCC TAC TAC TCA CAC GTG ATC CCC			4061
	Val Lys Leu Asp Lys Lys Asn Ala Tyr Tyr Ser His Val Ile Pro			
	1290	1295	1300	
35	AAG GAA GTC CTG AGC GAG GTC TTT GGG AAG GTT TTC CAG AAA AAC GTC			4109
	Lys Glu Val Leu Ser Glu Val Phe Gly Lys Val Phe Gln Lys Asn Val			
	1305	1310	1315	
40	ACT CCT CAG ACC TTC AGG AAG ATG GTC GAG GAC CGA AGA CTC GAT CCC			4157
	Ser Pro Gln Thr Phe Arg Lys Met Val Glu Asp Gly Arg Leu Asp Pro			
	1320	1325	1330	
45	GAA AAG GCC CAG AGG CTC TCC TGG CTC ATT GAG GGG GAC GTC GTG CTC			4205
	Glu Lys Ala Gln Arg Leu Ser Trp Leu Ile Glu Gly Asp Val Val Leu			
	1335	1340	1345	1350
50	GAC CGC GTT GAG TCC GTT GAT GTG GAA GAC TAC GAT GGT TAT GTC TAT			4253

	Asp Arg Val Glu Ser Val Asp Val Glu Asp Tyr Asp Gly Tyr Val Tyr			
5	1355	1360	1365	
	GAC CTG AGC GTC GAG GAC AAC GAG AAC TTC CTC GTT GGC TTT GGG TTG			4301
	Asp Leu Ser Val Glu Asp Asn Glu Asn Phe Leu Val Gly Phe Gly Leu			
10	1370	1375	1380	
	GTC TAT GCT CAC AAC AGC TAC TAC CCG TAC TAC GGC TAT GCA AGG GCG			4349
	Val Tyr Ala His Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr Ala Arg Ala			
15	1385	1390	1395	
	CGC TGG TAC TGC AAG GAG TGT GCA CAG AGC GTA ACG GCC TGG GGA AGG			4397
	Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser Val Thr Ala Trp Gly Arg			
20	1400	1405	1410	
	GAG TAC ATA ACG ATG ACC ATC AAG GAG ATA GAG GAA AAG TAC GGC TTT			4445
	Glu Tyr Ile Thr Met Thr Ile Lys Glu Ile Glu Glu Lys Tyr Gly Phe			
25	1415	1420	1425	1430
	AAG GTA ATC TAC AGC GAC ACC GAC GGA TTT TTT GCC ACA ATA CCT GGA			4493
	Lys Val Ile Tyr Ser Asp Thr Asp Gly Phe Phe Ala Thr Ile Pro Gly			
30	1435	1440	1445	
	GCC GAT GCT GAA ACC GTC AAA AAG AAG GCT ATG GAG TTC CTC AAC TAT			4541
	Ala Asp Ala Glu Thr Val Lys Lys Lys Ala Met Glu Phe Leu Asn Tyr			
35	1450	1455	1460	
	ATC AAC GCC AAA CTT CCG GGC GCG CTT GAG CTC GAG TAC GAG GGC TTC			4589
	Ile Asn Ala Lys Leu Pro Gly Ala Leu Glu Leu Glu Tyr Glu Gly Phe			
40	1465	1470	1475	
	TAC AAA CGC GGC TTC TTC GTC ACG AAG AAG AAG TAT GCG GTG ATA GAC			4637
	Tyr Lys Arg Gly Phe Phe Val Thr Lys Lys Lys Tyr Ala Val Ile Asp			
45	1480	1485	1490	
	GAG GAA GGC AAG ATA ACA ACG CGC GGA CTT GAG ATT GTG AGG CGT GAC			4685
	Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu Glu Ile Val Arg Arg Asp			
50	1495	1500	1505	1510

EP 0 745 675 A2

	TGG ACC GAG ATA GCG AAA GAG ACG CAG GCG AGG GTT CTT GAA GCT TTG	4733
5	Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala Arg Val Leu Glu Ala Leu	
	1515 1520 1525	
	CTA AAG GAC GGT GAC GTC GAG AAG GCC GTG AGG ATA GTC AAA GAA GTT	4781
10	Leu Lys Asp Gly Asp Val Glu Lys Ala Val Arg Ile Val Lys Glu Val	
	1530 1535 1540	
	ACC GAA AAG CTG AGC AAG TAC GAG GTT CCG CCG GAG AAG CTG GTG ATC	4829
15	Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro Pro Glu Lys Leu Val Ile	
	1545 1550 1555	
	CAC GAG CAG ATA ACG AGG GAT TTA AAG GAC TAC AAG GCA ACC GGT CCC	4877
20	His Glu Gln Ile Thr Arg Asp Leu Lys Asp Tyr Lys Ala Thr Gly Pro	
	1560 1565 1570	
	CAC GTT GCC GTT GCC AAG AGG TTG GCC GCG AGA GGA GTC AAA ATA CGC	4925
25	His Val Ala Val Ala Lys Arg Leu Ala Ala Arg Gly Val Lys Ile Arg	
	1575 1580 1585 1590	
	CCT GGA ACG GTG ATA AGC TAC ATC GTG CTC AAG GGC TCT GGG AGG ATA	4973
30	Pro Gly Thr Val Ile Ser Tyr Ile Val Leu Lys Gly Ser Gly Arg Ile	
	1595 1600 1605	
	GGC GAC AGG GCG ATA CCG TTC GAC GAG TTC GAC CCG ACG AAG CAC AAG	5021
35	Gly Asp Arg Ala Ile Pro Phe Asp Glu Phe Asp Pro Thr Lys His Lys	
	1610 1615 1620	
	TAC GAC GCC GAG TAC TAC ATT GAG AAC CAG GTT CTC CCA GCC GTT GAG	5069
40	Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln Val Leu Pro Ala Val Glu	
	1625 1630 1635	
	AGA ATT CTG AGA GCC TTC GGT TAC CGC AAG GAA GAC CTG CGC TAC CAG	5117
45	Arg Ile Leu Arg Ala Phe Gly Tyr Arg Lys Glu Asp Leu Arg Tyr Gln	
	1640 1645 1650	
	AAG ACC AGA CAG GTT GGT TTG AGT GCT TGG CTG AAG CCG AAG GGA ACT	5165
50	Lys Thr Arg Gln Val Gly Leu Ser Ala Trp Leu Lys Pro Lys Gly Thr	

55

1655	1660	1665	1670			
TGACCTTTCC	ATTTGTTTTC	CAGCCGATAA	CCCTTTAACT	TCCCTTC	AAACTCCCT	5225
5						
TAGGGAAAGA	CCATGAAGAT	AGAAATCCGG	CGGGCCCCGG	TTAAATACGC	TAGGATAGA	5285
GTGAAGCCAG	ACGGCAGGGT	ACTCGTCACT	GCCCCGACGG	TTCAACGTTG	AGAAGTT	5342

10 SEQ ID No.6

Length: 5342 base pairs

Type: nucleic acid (DNA)

15 Strandedness: double

Topology: linear

Molecular Type: cDNA

20 Source: hyperthermophilic archaeon

Strain name: KOD1

Sequence Description:

25	GCTTGACGGCC	CTGGCGTTAT	GGGACGTTGCC	AGTTTGC	GCC TACTCAAACA	TGCCGGTTTT	60
30	ATAAACGGAGA	AAAATGGGGA	GCTATTACGA	TCTCTCCTTG	ATGTGGGGTT	TACAATAAAG	120
	CCTGGATTGT	TCTACAAGAT	TATGGGGAT	GAAGATGAT	CCTCGACACT	GACTACATAA	
	CCGAGGATGG	AAAGCCTGTC	ATAAGAATT	TCAAGAAGGA	AAACGGCGAG	TTTAAGATTG	180
35	ACTACGACCG	GACTTTGAA	CCCTACTTCT	ACCCCTCCT	GAAGGACCAT	TCTCCCATTG	240
	AGGAAGTCAA	GAAGATAACC	GCCGAGAGGC	ACGGGACGGT	TGTAACGGTT	AAGGGGGTTG	
	AAAAGGTMCA	GAAGAAGTTC	CTCCGGAGAC	CACTTGAGGT	CTGGAAACTC	TACTTTACTC	360
40	ATCCCGAGGA	CCTCCCAGCG	ATAAGGGACA	AGATACGAGA	GCATGGAGCA	GTTATTGACA	420
	TCTACGAGTA	CGACATACCC	TTGCCAAGC	GCTACCTCAT	AGACAAAGGA	TTAGTGCCAA	
	TGGAAGGCCA	CGAGGAGCTG	AAAATGCTCG	CCTTCGACAT	TCAAACCTCTC	TACCATGAGG	540
45	GCGAGGACTT	CCCCGACGGG	CCAATCCTTA	TGATAAGCTA	CGCCGACCGAG	GAAGGGGCCA	600
	GGCTGATAAC	TTGGAAGAAC	GTGGATCTCC	CCTACCTTGA	CGTCGTCTCG	ACGGAGAGGG	
	AGATGATAAA	CCGCTTCCTC	CGTGTGTGA	AGGAGAAAGA	CCCGGACGTT	CTCATAACCT	720
50	ACAACGGCGA	CAACTTCGAC	TTCGCCTATC	TGAAAAGCG	CTGTAAAAAG	CTCGGAATAA	780
	ACTTCGCCCT	CGGAAGGGAT	GGAAAGCGAGC	CGAACATTCA	GAGGATGGGC	GACAGGTTTG	
							840
							900

55

5 CCGTCGAAGT GAAGGGACGG ATACACTTCG ATCTCTATCC TGTGATAACA CGGACCATAA 960
 ACCTGCCAC ATACACCGTT GACCCCGTT ATGAAGCCGT CTTCGGTAG CGGAAGGAGA 1020
 AGCTTACGC TGAAGAAATA ACACCAGCCT GGGAAACCGG CGAGAACCTT GAGAGACTCG 1080
 CCCGCTACTC CATGGAACAT CGGAAGGTCA CATAACCACCT TGGGAAGGAG TTCCCTCCGA 1140
 10 TGGAGGCCCA CCTTCTCGC TTAATCGGCC AGTCCCTCTG GGACGTCTCC CGCTCCAGCA 1200
 CTGGCAACCT CGTTGAGTGG TTCCCTCTCA GGAAGGCCCT ATGAGAGGAA TGAGCTGGCC 1260
 CCGAACAAAG CCGATGAAA CGACGCTGGCC AGAAGACGCC AGAGCTATGA AGGAGGCTAT 1320
 15 GTAAAAGAGC CCGAGACAGG CTTGTCGGAG AACATAGTGT ACCTAGATTT TAGATGCCAT 1380
 CCAGCCGATA CGAAGGTGT CCTCAAGGGG AACGGGATTA TAAACATCAG CGAGGTTCA 1440
 GAAGGTGACT ATGTCCTTGG GATTGACGGC TGGCAGAGAG TTAGAAAAGT ATGGAATAC 1500
 20 GACTACAAAG GGGAGCTTGT AAACATAAAC CGCTTAAAGT CTACGCCAA TCATAAGCTT 1560
 CCCGTTGTTA CAAAGAACGA ACCACAAACG AGAATAAGAG ACAGTCTTGC TAAGTCTTTC 1620
 CTTACTAAAA AAGTTAAGGG CAACATAATA ACCACTCCCC TTTCTATGA AATAGGCAGA, 1680
 25 GCGACAAAGTC AGAATATTCC AGAAGAACAG GTTCTCAAGG GAGACCTCGC TGGCATAGTA, 1740
 TTGGCTGAAG GAACGCTCTT GAGCAAACAC GTTCAATACT TTGATTCACT CCCAAAAAAA 1800
 CCGAGGATT T CACACCAGTA TCGTCTTGAG ATAACCATTG CGAAAGACGA CGAGGAGTTT 1860
 30 AGGGATCGTA TCACATACAT TTTGAGCGT TTGTTTGGGA TTACTCCAAG CATCTCGGAG 1920
 AAGAAAGGAA CTAACCGAGT AACACTCAA GTTGCAGAAGA AGAATGTTA TCTTAAAGTC 1980
 AACCAAATT A TGGACAAACAT AGAGTCCCTA CATGCCCT CGTTCTCA 2040
 35 GAAGGCAGG GTTCAGTAAA CAGCTTAGGA GGAGTATTGT TGCACCCAG GGTACAAAGA 2100
 ACGAGTGGAA GATTAACCTG GTGTCAAAAC TGCTCTCCCA CCTTGGTATC CCTCATCAA 2160
 40 CGTACACGTA TCACTATCAG GAAAATGGGA AAGATCGGAG CAGGTATATA CTGGAGATAA 2220
 CTGGAAAGGA CGGATTGATA CTGTTCCAAA CACTCATTGG ATTCACTAGT GAAAGAAAGA 2280
 ACGCTCTGCT TAATAAGGCA ATATCTAGA GGGAAATGAA CAACTGGAA AACAAATGGAT 2340
 45 TTTACAGGCT CACTCAATTG AATGTCAGCA CGGAATACTA TGAGGGCAAG GTCTATGACT 2400
 TAACTCTTGA AGGAACCTCCC TACTTTGCCA ATGGCATATT GACCCATAAC TCCCTGTACC 2460
 CCTCAATCAT CATCACCCAC AACGTCTCGC CGGATACGCT CAACAGAGAA CGATGCAAGG 2520
 50 AATATCACGT TGGCCCCACAG GTCGGCCACC CCTTCTGCAA CGACTTCCCA GGATTTATCC 2580
 CGAGCCTGCT TGGAGACCTC CTAGAGGAGA CGCAGAAGAT AAAGAAGAAG ATGAAGGCCA 2640

5

CGATTGACCC GATCGAGAGG AACCTCCTCG ATTACAGGA GAGGGCCATC AAGATCCTGG 2700
 CAAACAGCAT CCTACCCGAC GAATCCCTTC CAGTCCTCGA GGAAGGGGAG GTTCACCTTCG 2760
 TCAGGATTGG AGAGCTCATA GACCCGATGA TGGAGGAAAA TGCTGGGAAA CTAAAGAGAG 2820
 AGGGCGAGAC GGAAGTGCTT CACGTCACTG GGCTTGAAGT CCCGTCTTT AACAGGAGAA 2880
 CTAACAAGGC CGAGCTCAAG AGACTAAAGG CCCTGATTAG CCACGATTAT TCTGGCAAGG 2940
 TCTACACCAT CACACTGAAG TCGGGGAGGA GAATAAACAT AACCTCTGGC CACAGCCTCT 3000
 TCTCTGTGAG AAACGGGGAG CTCCGTTGAAG TTACGGGGCA TGAACTAAAG CCAGGTGACC 3060
 TCGTTGCAGT CCCGCGGAGA TTGGACCTTC CTGAGAGAAA CCACGTGCTG AACCTCGTTG 3120
 AACTGCTCCT TGGAACCCCCA GAAGAAGAAA CTTTGGACAT CGTCATGACG ATCCCAGTCA 3180
 AGGTAAGAA GAACTTCTTT AAAGGGATGC TCAGGACTTT GCGCTGGATT TTCGGAGAGG 3240
 20 AAAACAGGCC CAGAACCCCG AGACGGTATC TCAGGCACCT TCAGGATCTG GGCTATGTCC 3300
 GGCTTAAGAA GATCGGCTAC GAAGTCCTCG ACTGGGACTC ACTTAAGAAC TACAGAAGGC 3360
 TCTACGAGGC GCTTGTCCAG AACGTCAAGAT ACAACGGCAA CAAGAGGGAG TACCTCGTTG 3420
 25 AATTCAATTTC CATCCGGGAT GCAGTTGGCA TAATGCCCT AAAAGAGCTG AAGGAGTGGA 3480
 AGATCGGCAC CCTGAACGGC TTCACAATGA CAAAGCTCAT TGAAGTGGAC GAGTCGTTAG 3540
 CAAAGCTCCT CGGCTACTAC GTGACCGAGG GCTATCCAAG AAAGCAGAGG AATCCAAAAA 3600
 30 ACGGCTGGAG CTACAGCGTG AAGCTCTACA ACCAAGACCC TGAAGTGGCTG GACGATAATGG 3660
 AGAGACTCCC CAGCAGGTTT TTGGGGAGG TGAGGCGGGG CAGGAACTAC GTTGAGATAC 3720
 CGAACAAGAT CGGCTACCTG CTCTTGAGA ACATGTGGG TGTCTCTAGG GAGAACAAAGA 3780
 35 GGATTCGGGA GTTCGTCTTC ACGTCCCCGA AAGGGGTMCG, GCTGGCCTTC CTTCAGGGGT 3840
 ACTCATCGGC GATGGCGACG TCCACCGAAC AAGAGACTCA GGCTCTCAAC GAAAAGCGAG 3900
 40 CTTCAGCGAA CCAGCTCGTC CTCCCTTTGA ACTCGGTGGG GGTCTCTGCT GTAAAACCTTG 3960
 GGCACCGACAG CGGGCTTTAC AGGCTCTATA TAAACCGAGGA GCTCCGTTTC GTAAAGCTGG 4020
 ACAAGAAAAA GAACGCCAAC TACTCACACG TGATCCCCAA GGAAGTCCTG AGCGAGGTCT 4080
 45 TTGGGAAGGT TTTCAGAAA AACGTCACTC CTCAGACCTT CAGGAACATG CTCGAGGACG 4140
 GAAGACTCGA TCCCGAAAAG GCCCAGAGGC TCTCCCTGGCT CATTGAGGGG GACGTAGTGC 4200
 TCGACCGCGT TGAGTCCGTT CATGTGGAG ACTACCGATGG TTATGTCTAT GACCTGAGCG 4260
 50 TCGAGGACAA CGAGAACTTC CTCCGTTGGCT TTGGGTTGGT CTATGCTCAC AACAGCTACT 4320
 ACGGTTACTA CGGCTATGCA AGGGCCCGCT GGTACTCCAA GGACTGTGCA GAGAGCGTAA 4380

55

CGGCCCTGGGG AAGGGAGTAC ATAACGATGA CCATCAAGGA CATAAGGGAA AAGTACGGCT 4440
5 TTAAGGTAAT CTACAGCGAC ACCGACGGAT TTTTGCCAC AATACTGGG 4500
AAACCGTCAA AAAGAAGGCT ATGGAGTTCC TCAACTATAT CAACGCCAAA CTTCCGGGG 4560
CGCTTCAGCT CGAGTACCGAG GGCTTCTACA AACCGGGCTT CTTCGTCACG AAGAAGAACT 4620
10 ATGCGGTGAT AGACGAGGAA GGCAAGATAA CAACGCCGG ACTTGAGATT GTGAGCCGTG 4680
ACTGGAGCGA GATAGCGAAA GAGACCGAGG CGAGGTTCT TGAAGCTTG CTAAGGACG 4740
15 GTGACGTCGA GAAGGCCGTG AGGATAGTCA AAGAAGTTAC CGAAAAGCTG AGCAAGTACG 4800
AGGTTCCGCC GGAGAAGCTG GTGATCCACG ACCAGATAAC GAGGGATTTA AAGGACTACA 4860
AGGCAACCGG TCCCCACGTT CCCGTTGCCA AGAGGTTGGC CGCGAGAGGA GTCAAAATAC 4920
20 GCCCTCGAAC CGTGATAACC TACATCGTGC TCAAGGGCTC TGGGAGGATA GGCGACAGGG 4980
CGATACCGTT CGACCGAGTTC GACCCGACGA ACCACAAGTA CGATGCCGAG TACTACATTG 5040
25 AGAACCCAGGT TCTCCAGCC GTTGAGACAA TTCTGAGAGC CTTCGGTTAC CGCAAGGAAG 5100
ACCTGCGCTA CCAGAAGACG AGACAGGTTG GTTTGAGTGC TTGGCTGAAG CGGAAGGGAA 5160
30 CTTGACCTTT CCATTTGTT TCCACCGGAT AACCCCTTAA CTTCCCTTTC AAAAACTCCC 5220
TTTAGGGAAA GACCATGAAG ATAGAAATCC GGGGGGGCCC GGTAAATAC GCTAGGATAG 5280
AACTGAAGCC AGACGGCAGG GTACTCGTCA CTGCCCCGAG GGTCAACGT TGAGAAGTT 5339
35

40

45

50

55

SEQ ID No.7

Length: 24 base pairs

5 Type: nucleic acid

Topology: linear

10 Molecular Type: synthetic DNA

Sequence Description:

GGATTAGTGC CAATGGAAGG CGAC 24

15 SEQ ID No.8

Length: 24 base pairs

Type: nucleic acid

20 Topology: linear

Molecular Type: synthetic DNA

Sequence Description:

25 GAGGGCGAAG TTTATTCCGA GCTT 24

SEQ ID No.9

Length: 324 base pairs

30 Type: nucleic acid (DNA)

Strandedness: double

Topology: linear

35 Molecular Type: cDNA

Sequence Description:

40 GGATTAGTGC CAATGGAAGG CGACGAGGAG CTGAAAATGC TCCCTTCGA CATTCAA 60

CTCTACCATG AGGGCGAGGA GTTCCGGAG GGGCCAATCC TTATGATAAG CTACGCCGAC 120

GAGGAAGGGG CCAGGGTGTAT AACTTGGAAG AACGTGGATC TCCCCTACGT TGACGTCGTC 180

45 TCGACGGAGA GGGAGATGTAT AAAGCGCTTC CTCCGTGTTG TGAAGGAGAA AGACCCGGAC 240

GTTCTCATAA CCTACAAACCG CGACAACTTC GACTTCGCCT ATCTGAAGAA CGGCTGTGAA 300

AAGCTCGGAA TAAACTTCCC CCTC 324

50

55

SEQ ID No.10

Length: 108 base pairs

Type: amino acid

Topology: linear

Molecular Type: protein

Sequence Description:

Gly Leu Val Pro Met Glu Glu Gly Asp Glu Glu Leu Lys Met Leu Ala Phe

1 5 10 15

Asp Ile Gin Thr Leu Tyr His Glu Gly Glu Glu Phe Ala Glu Gly Pro

20 25 30

Ile Leu Met Ile Ser Tyr Ala Asp Glu Glu Gly Ala Arg Val Ile Thr

35 40 45

Trp Lys Asn Val Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Arg

50 55 60

Glu Met Ile Lys Arg Phe Leu Arg Val Val Lys Glu Lys Asp Pro Asp

65 70 75 80

Val Leu Ile Thr Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys

85 90 95

Lys Arg Cys Glu Lys Leu Gly Ile Asn Phe Ala Leu

100 105

45

50

55

55
SEQ ID No.11

Length: 42 base pairs

5
Type: nucleic acid (DNA)

Strandedness: single

10
Molecular Type: synthetic DNA

Sequence Description:

GCCATCAAGA TCCTGGCAAA CAGCTACTAC GGTTACTACG GC 42

15
SEQ ID No.12

Length: 32 base pairs

20
Type: nucleic acid (DNA)

Strandedness: single

25
Molecular Type: synthetic DNA

Sequence Description:

25
GATGGATCCA ACTTCTCAAC GTTGAACCCT CG 32

SEQ ID No.13

30
Length: 46 base pairs

Type: nucleic acid (DNA)

35
Strandedness: single

Molecular Type: synthetic DNA

Sequence Description:

40
GAACATAGTG TACCTAGATT TTAGATCCCT GTACCCCTCA ATCATC 46

SEQ ID No.14

Length: 42 base pairs

45
Type: nucleic acid (DNA)

Strandedness: single

50
Molecular Type: synthetic DNA

Sequence Description:

55
GCCGTAGTAA CCGTAGTAGC TGTTGCCAG GATCTTGATG GC 42

SEQ ID No.15

Length: 33 base pairs

Type: nucleic acid (DNA)

Strandedness: single

Molecular Type: synthetic DNA

Sequence Description:

ATCGATATCC TCGACACTGA CTACATAACC GAG

33

SEQ ID No.16

Length: 46 base pairs

Type: nucleic acid (DNA)

Strandedness: single

Molecular Type: synthetic DNA

Sequence Description:

GATGATTGAG GGGTACAGGG ATCTAAAATC TAGGTACACT ATGTTTC

46

Claims

1. A DNA comprising a nucleotide sequence that encodes a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity derived from a hyperthermophilic archaeon strain.

2. The DNA of claim 1, wherein the strain is KOD1.

3. The DNA of claim 1 or 2, wherein a nucleotide sequence encodes the thermostable DNA polymerase which has the amino acid sequence of SEQ ID No. 1.

4. The DNA of any one of claims 1 to 3, wherein said DNA has a nucleotide sequence of SEQ ID No. 5.

5. A DNA encoding a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity, said DNA comprising a nucleotide sequence:

(a) which differs from a DNA of claim 4 in codon sequence due to the degeneracy of the genetic code;

(b) which hybridizes with a DNA of claim 4 or section (a), above; or

(c) represents a fragment, allelic or other variation of a DNA of claim 4.

6. A recombinant DNA expression vector that comprises the DNA sequence of any one of claims 1 to 5.

7. The recombinant DNA expression vector of claim 6, in which the vector is a vector derived from pET-8c.

8. A recombinant host cell which is transformed by a recombinant DNA expression vector of claim 7.

9. The recombinant host cell of claim 8, in which the host cell is Escherichia coli.

10. A thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity which is encoded by the DNA of any one of claims 1 to 5.

11. The thermostable DNA polymerase of claim 10, which has the following physical and chemical properties:
5 Action: It catalyses the extension reaction of nucleotide sequence that is complementary to a template nucleotide sequence using nucleotide triphosphates as substrate and it has a 3'-5' exonuclease activity,
DNA extension rate: at least 30 bases/second
Optimum pH: 6.5 - 7.5 (at 75°C)
Optimum temperature: 75°C
10 Molecular weight: about 88 - 90 kDa
Amino acid sequence: as defined in SEQ ID No. 1.

12. A method for producing a DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity comprising culturing the recombinant host cell of claim 8 or 9.
15

13. A method for purifying a DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity comprising culturing the recombinant host cell of claim 8 or 9, and further (a) recovering the cultured recombinant host cells, disintegrating them and preparing the cell extract, and (b) removing the impurified proteins derived from recombinant host cells.
20

14. The DNA polymerase of claim 10 or 11, which is obtainable by using the recombinant host cells of claim 8 or 9 or by the method of claims 12 or 13.
25

15. A method for amplifying a target nucleic acid comprising reacting the target nucleic acid with four kinds of dNTP and primer sufficiently complementary to said target nucleic acid in a buffer solution which contains a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity such that the above mentioned primer is annealed to the target nucleic acid and an extention product is synthesized from the primer.
30

16. A method for amplifying a target nucleic acid in a sample wherein each target nucleic acid consists of two separate complementary strands which comprises the following steps (a) to (d), characterized in that a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity is used as a thermostable DNA polymerase;
35
(a) modifying the target nucleic acid, if necessary, to produce single-stranded nucleic acids;
(b) reacting the single-stranded nucleic acids with four kinds of dNTP and primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, in a buffer solution which contains said thermostable DNA polymerase such that the above mentioned primers are annealed to the single-stranded nucleic acids and extention products are synthesized from the primers;
40
(c) separating the primer extention products from the templates on which they are synthesized to produce single-stranded nucleic acids; and
(d) repeatedly conducting the above mentioned steps (b) and (c).
45

17. A method for detecting a target nucleic acid in a sample wherein each target nucleic acid consists of two separate complementary strands which comprises the following steps (a) to (e), characterized in that a thermostable DNA polymerase having a DNA extension rate of at least 30 bases/second and a 3'-5' exonuclease activity is used as a thermostable DNA polymerase:
50
(a) modifying the target nucleic acid, if necessary, to produce single-stranded nucleic acids;
(b) reacting the single-stranded nucleic acids with four kinds of dNTP and primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, in a buffer solution which contains said thermostable DNA polymerase such that the above mentioned primers are annealed to the single-stranded nucleic acids and extention products are synthesized from the primers;
(c) separating the primer extention products from the templates on which they are synthesized to produce single-stranded nucleic acids;
55
(d) repeatedly conducting the above mentioned steps (b) and (c); and
(e) detecting an amplified nucleic acid.
56

18. The method of any one of claims 15 to 17, wherein said thermostable DNA polymerase is encoded by the DNA of

EP 0 745 675 A2

any one of claims 1 to 5 or produced by the method of claim 12 or 13.

5 19. A reagent kit for amplifying target nucleic acid which comprises primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, four kinds of dNTP, divalent cation and the thermostable DNA polymerase encoded by the DNA of any one of claims 1 to 5 or produced by the method of claim 12 or 13.

10 20. A reagent kit for detecting target nucleic acid which comprises primers, wherein said primers are selected so as to be sufficiently complementary to different strands of target nucleic acid to anneal therewith, four kinds of dNTP, divalent cation, thermostable DNA polymerase encoded by the DNA of any one of claims 1 to 5 or produced by the method of claim 12 or 13, amplifying buffer solution, a probe capable of hybridizing with amplified nucleic acid and a detection buffer solution.

15

20

25

30

35

40

45

50

55

Fig. 1

Measurement of the DNA Extension Rate of DNA Polymerase
Derived from the Hyperthermophilic Archaeon Strain KOD1

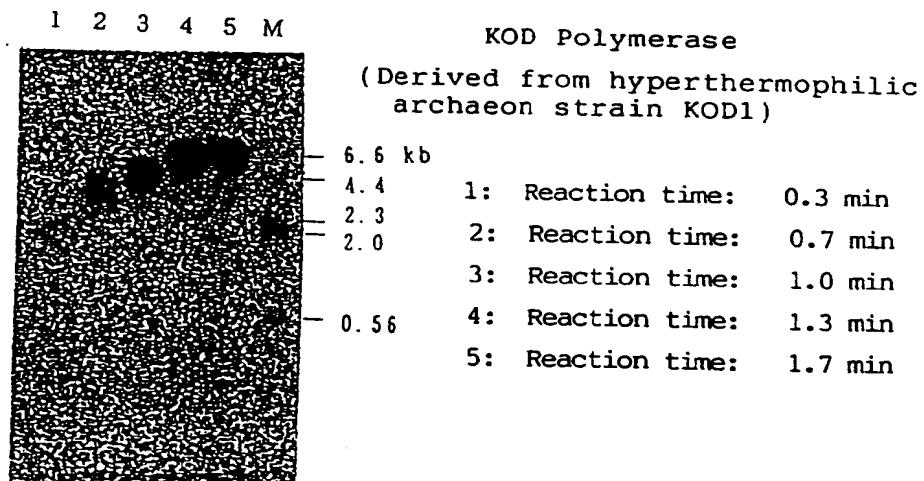
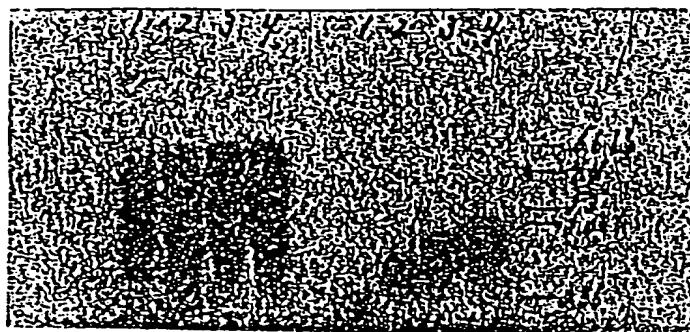
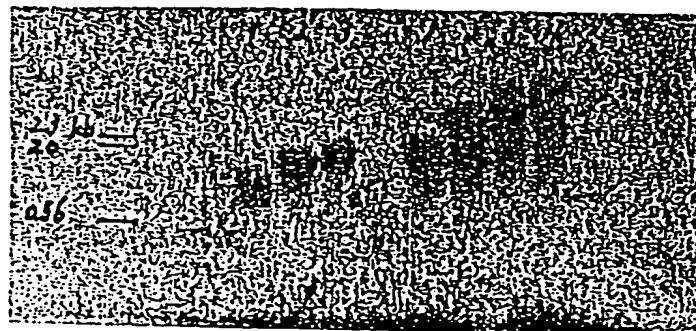


Fig. 2

Comparison of DNA Extension Rates of Various Thermostable DNA Polymerases

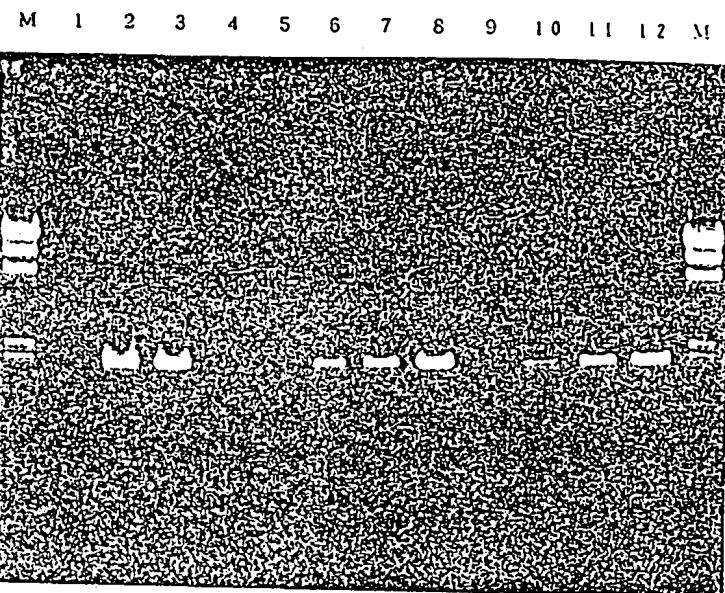


	KOD Polymerase	Pfu Polymerase
1:	Reaction Time: 0.7 min	1: Reaction Time: 1.0 min
2:	" : 1.0 min	2: " : 1.5 min
3:	" : 1.3 min	3: " : 2.0 min
4:	" : 1.7 min	4: " : 2.5 min



	Deep Vent Polymerase	Taq Polymerase
1:	Reaction Time: 1.5 min	1: Reaction Time: 1.0 min
2:	" : 2.0 min	2: " : 1.5 min
3:	" : 2.5 min	3: " : 2.0 min
4:	" : 2.5 min	4: " : 2.5 min

Fig. 3

Comparison of the PCR due to the Difference in the Reaction Time
of Various Thermostable DNA Polymerases

- 1: KOD Polymerase (Derived from Hyperthermophilic archaeon strain KOD1) Reaction Time: 1 sec.
- 2: KOD Polymerase (Derived from Hyperthermophilic archaeon strain KOD1) " " : 5 sec.
- 3: KOD Polymerase (Derived from Hyperthermophilic archaeon strain KOD1) " " : 10 sec
- 4: Taq Polymerase (Derived from *Thermus aquaticus*) Reaction Time: 1 sec.
- 5: Taq Polymerase (Derived from *Thermus aquaticus*) " " : 5 sec.
- 6: Taq Polymerase (Derived from *Thermus aquaticus*) " " : 10 sec.
- 7: Taq Polymerase (Derived from *Thermus aquaticus*) " " : 30 sec.
- 8: Taq Polymerase (Derived from *Thermus aquaticus*) " " : 60 sec.
- 9: Pfu Polymerase (Derived from *Pyrococcus furiosus*) " " : 30 sec.
- 10: Pfu Polymerase (Derived from *Pyrococcus furiosus*) " " : 60 sec.
- 11: Pfu Polymerase (Derived from *Pyrococcus furiosus*) " " : 90 sec.
- 12: Pfu Polymerase (Derived from *Pyrococcus furiosus*) " " : 120 sec.

Fig. 4

Construction of Expression Recombinant Vector (pET-pol)

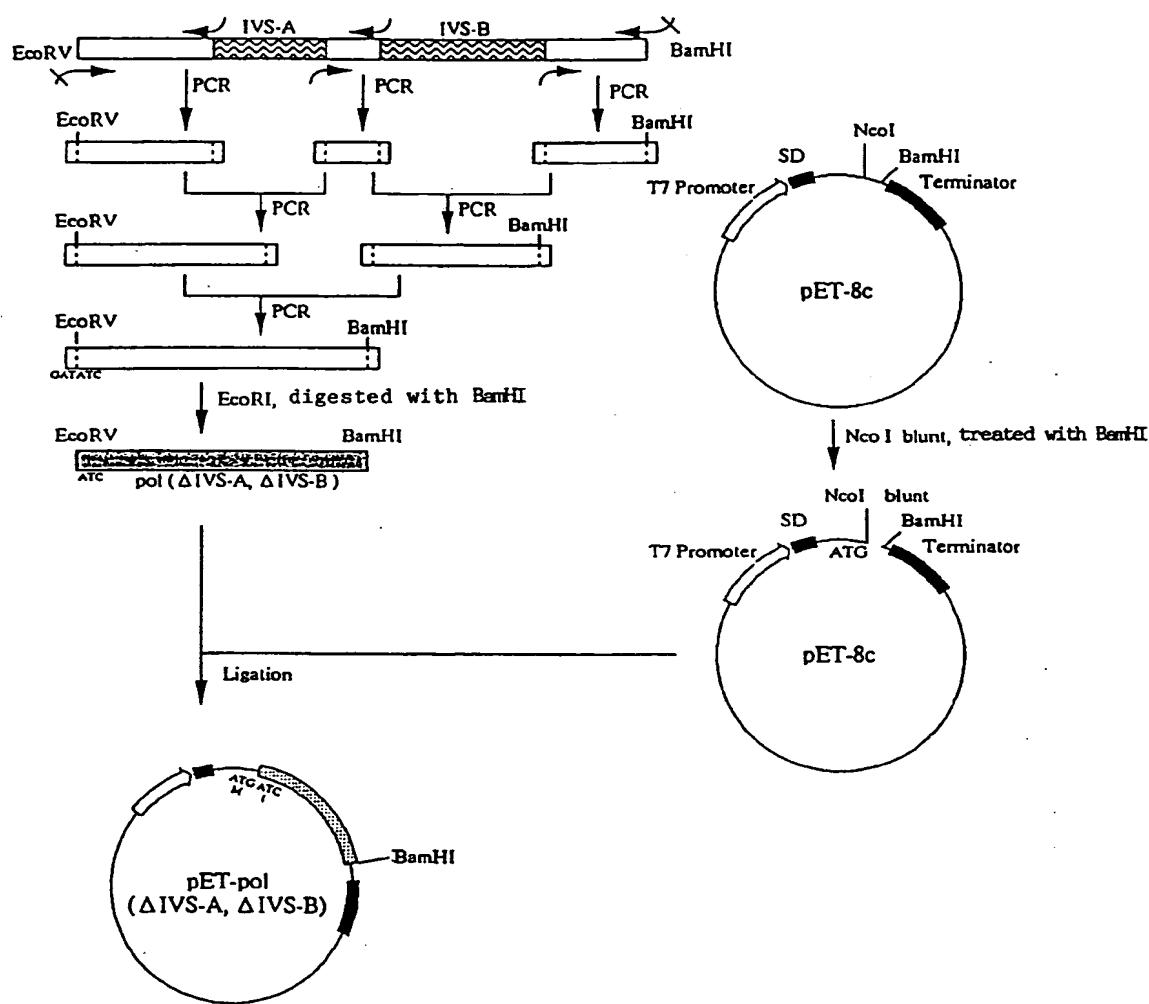
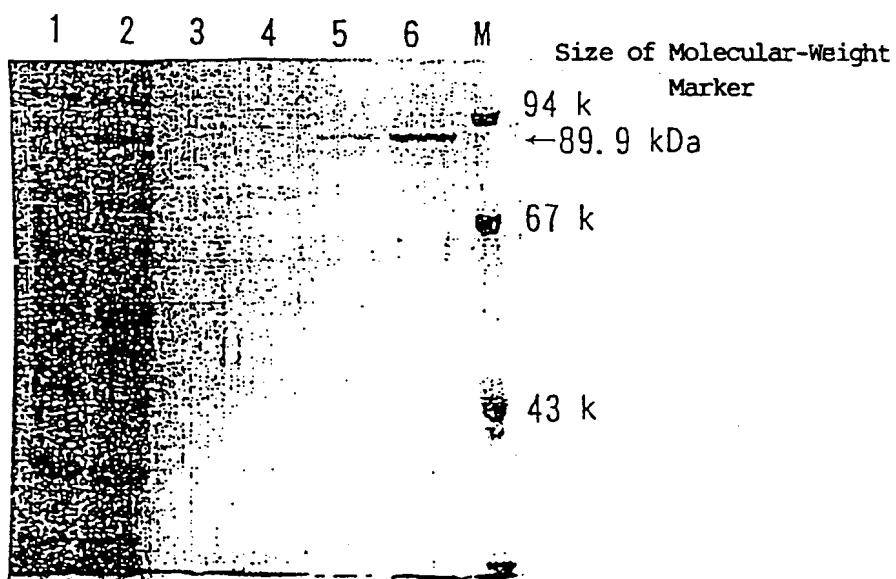
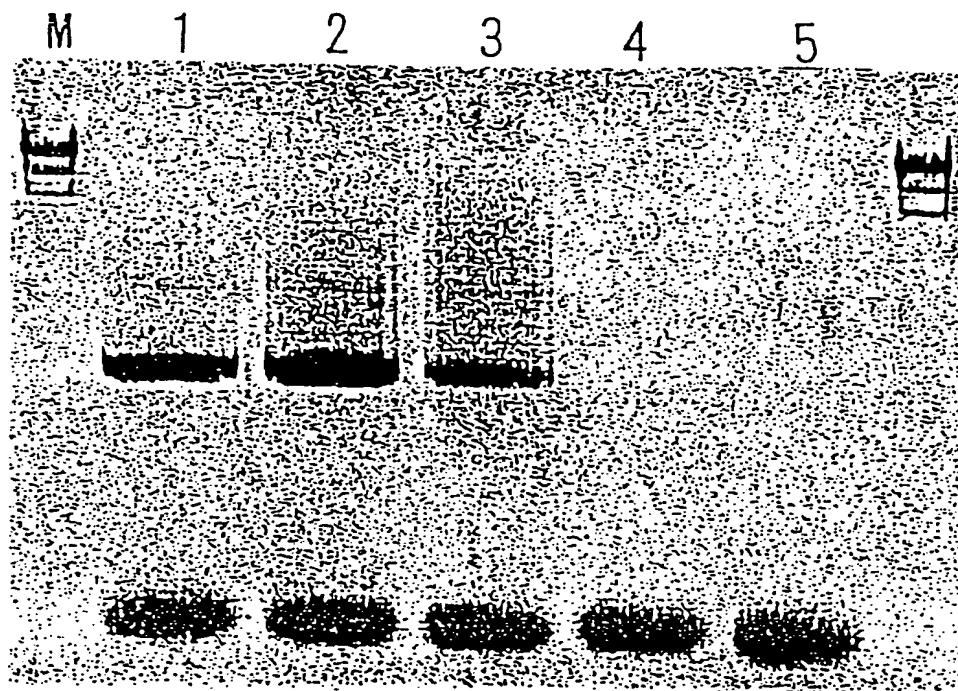


Fig. 5



- 1: pET-8c Precipitate
- 2: pET-pol (Δ IIVS-A, Δ IIVS-B) Precipitate
- 3: pET-8c Supernatant Liquid
- 4: pET-8c Supernatant Liquid \times 5
- 5: pET-pol (Δ IIVS-A, Δ IIVS-B) Supernatant Liquid
- 6: pET-pol (Δ IIVS-A, Δ IIVS-B) Supernatant Liquid \times 5

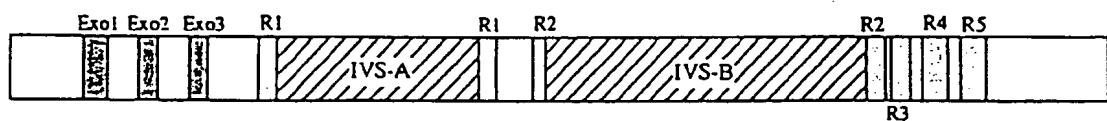
Fig. 6



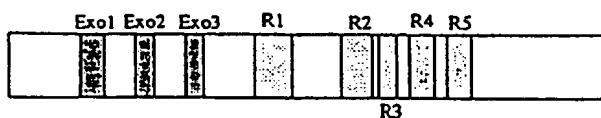
- 1: Vent Polymerase (Derived from *Thermococcus litoralis*)
- 2: pET-pol (Δ IIVS-A, Δ IIVS-B) Supernatant Liquid
- 3: pET-pol (Δ IIVS-A, Δ IIVS-B) Supernatant Liquid \times 5
- 4: pET-8c Supernatant Liquid
- 5: pET-8c Supernatant Liquid \times 5

Fig. 7

DNA Polymerase Gene of Hyperthermophilic Archaeon Strain KOD1



DNA Polymerase Gene of Pyrococcus furiosus (Pfu DNA Polymerase)



DNA Polymerase Gene of Thermococcus litoralis (Vent DNA Polymerase)

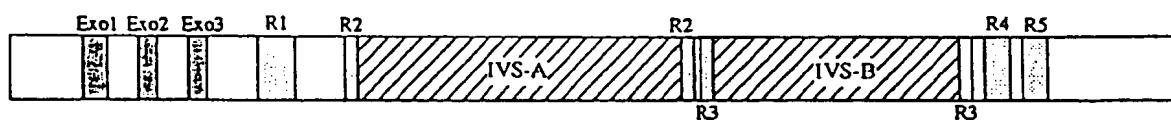


Fig. 1

Measurement of the DNA Extension Rate of DNA Polymerase
Derived from the Hyperthermophilic Archaeon Strain KOD1

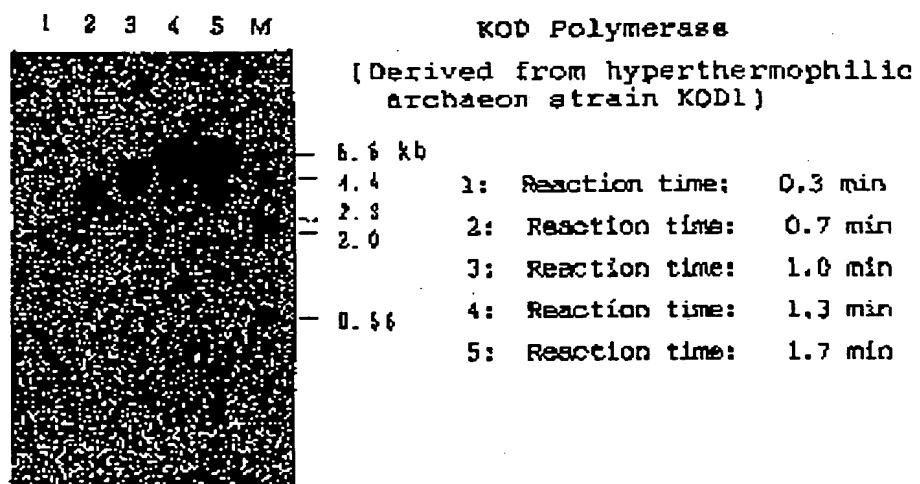
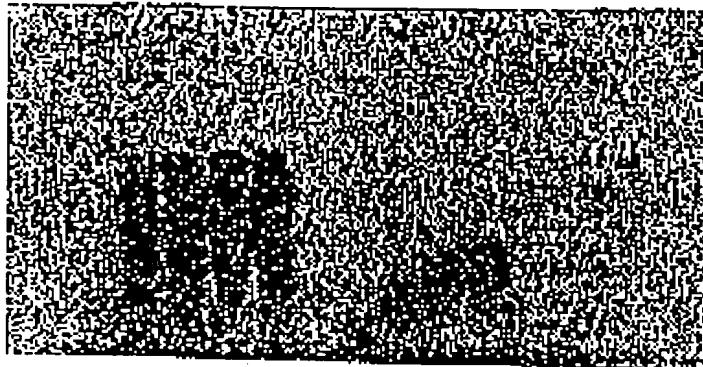


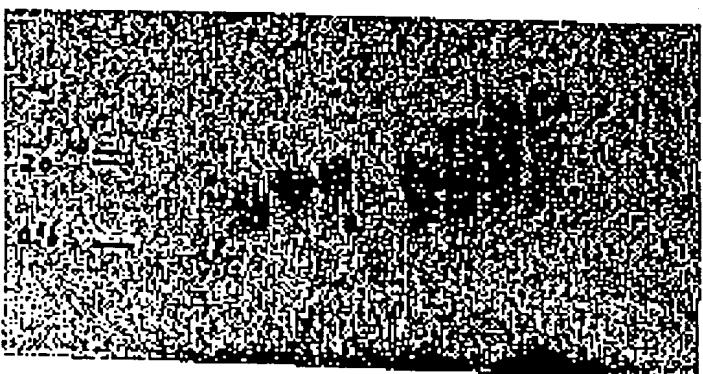
Fig. 2

Comparison of DNA Extension Rates of Various Thermostable
DNA Polymerases



2 a

KOD Polymerase		Pfu Polymerase	
1:	Reaction Time: 0.7 min	1:	Reaction Time: 1.0 min
2:	" " : 1.0 min	2:	" " : 1.3 min
3:	" " : 1.3 min	3:	" " : 1.6 min
4:	" " : 1.7 min	4:	" " : 2.0 min

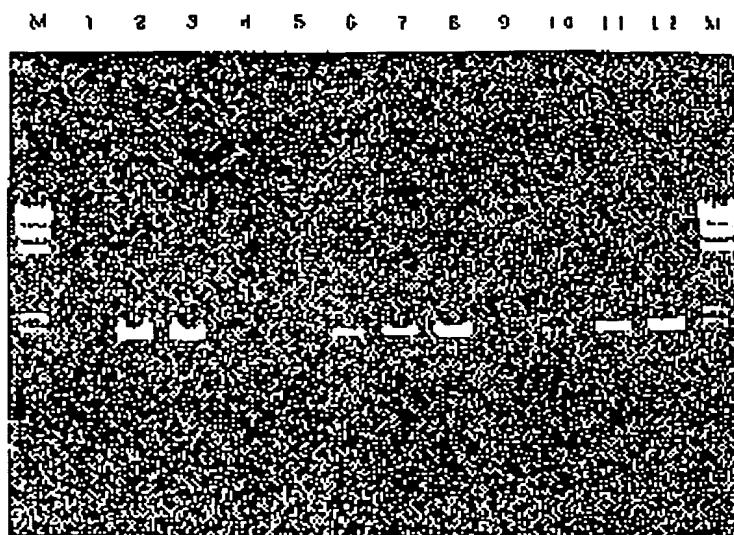


2 b

Deep Vent Polymerase		Taq Polymerase	
1:	Reaction Time: 1.5 min	1:	Reaction Time: 1.0 min
2:	" " : 2.0 min	2:	" " : 1.5 min
3:	" " : 2.5 min	3:	" " : 2.0 min
		4:	" " : 2.5 min

Fig. 3

Comparison of the PCR due to the Difference in the Reaction Time of Various Thermostable DNA Polymerases



1: KOD Polymerase (Derived from Hyperthermophilic archaeon strain KOD1)	Reaction Time: 1 sec.
2: KOD Polymerase (Derived from Hyperthermophilic archaeon strain KOD1)	" : 5 sec.
3: KOD Polymerase (Derived from Hyperthermophilic archaeon strain KOD1)	" : 10 sec.
4: Taq Polymerase (Derived from <i>Thermus aquaticus</i>)	Reaction Time: 1 sec.
5: Taq Polymerase (Derived from <i>Thermus aquaticus</i>)	" : 5 sec.
6: Taq Polymerase (Derived from <i>Thermus aquaticus</i>)	" : 10 sec.
7: Taq Polymerase (Derived from <i>Thermus aquaticus</i>)	" : 30 sec.
8: Taq Polymerase (Derived from <i>Thermus aquaticus</i>)	" : 60 sec.
9: Pfu Polymerase (Derived from <i>Pyrococcus furiosus</i>)	" : 30 sec.
10: Pfu Polymerase (Derived from <i>Pyrococcus furiosus</i>)	" : 60 sec.
11: Pfu Polymerase (Derived from <i>Pyrococcus furiosus</i>)	" : 90 sec.
12: Pfu Polymerase (Derived from <i>Pyrococcus furiosus</i>)	" : 120 sec.

Fig. 4

Construction of Expression Recombinant Vector (pET-pol)

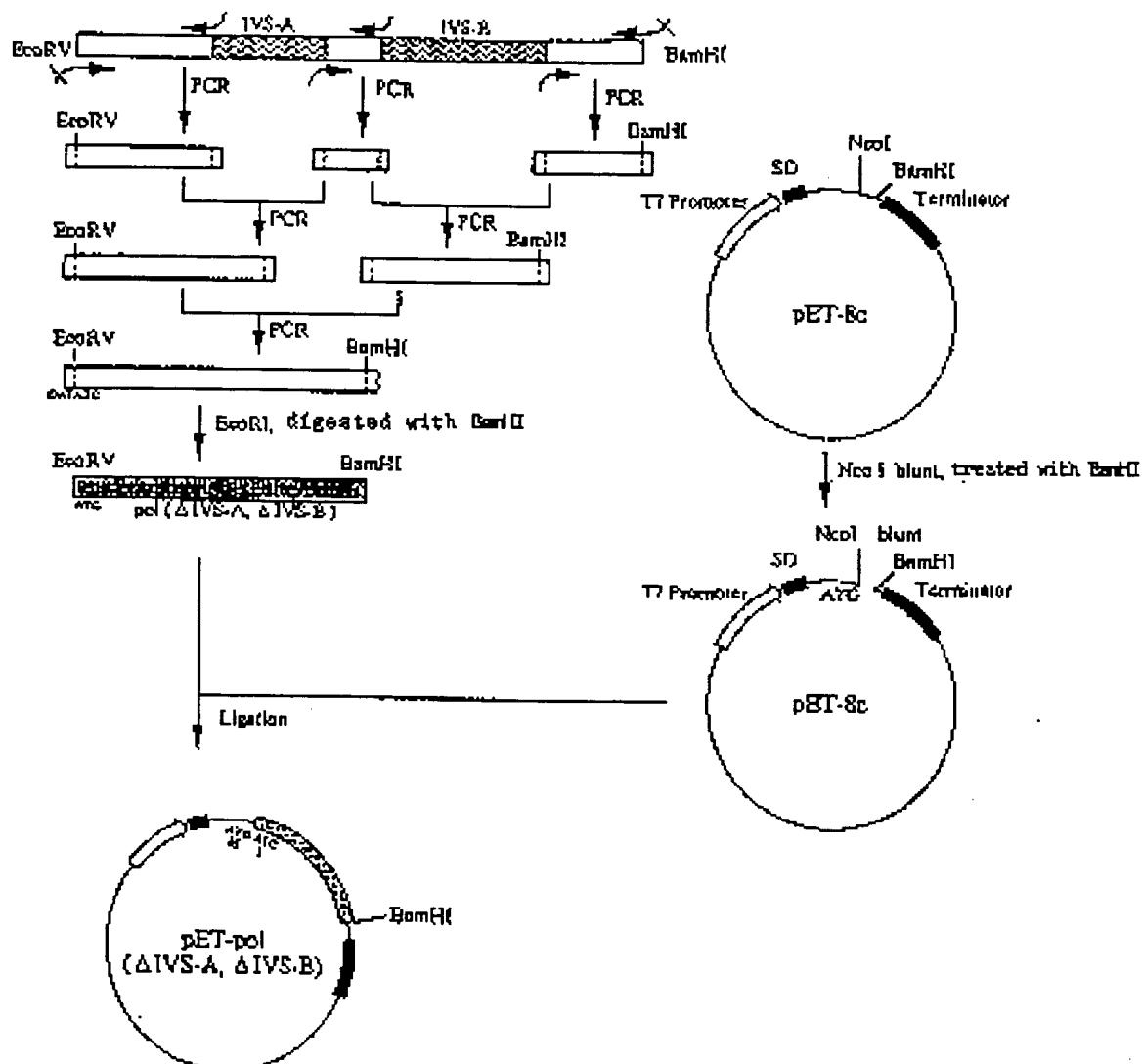
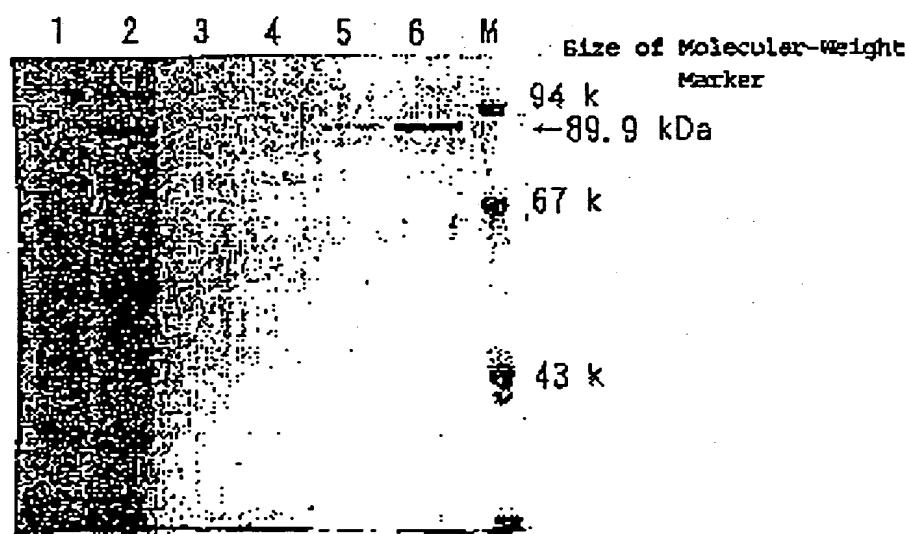
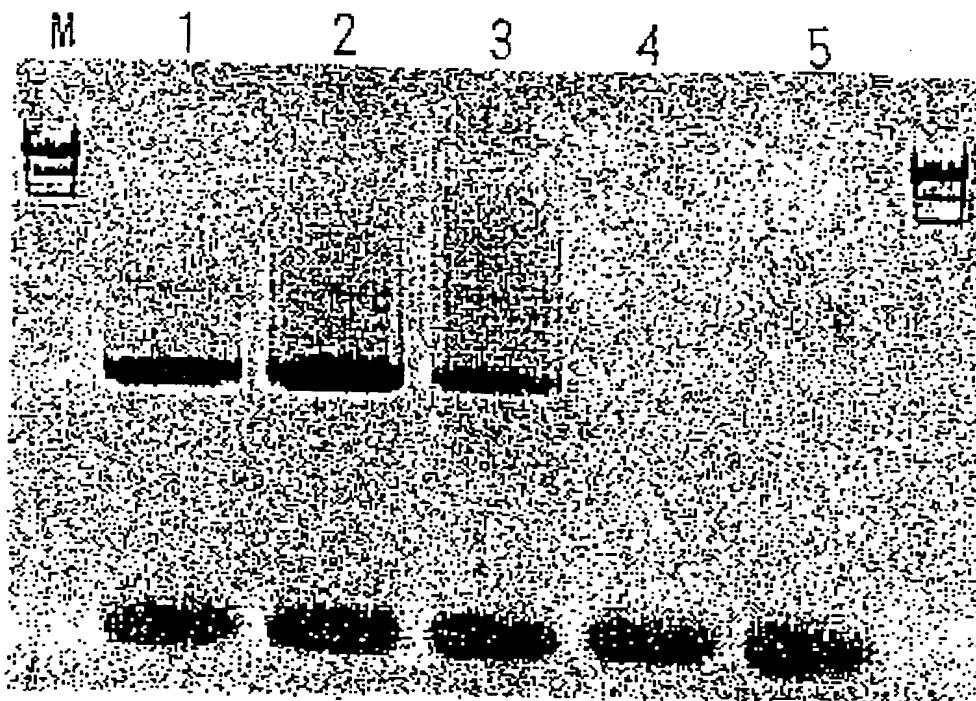


Fig. 5



- 1: pET-8c Precipitate
- 2: pET-pol (ΔIVS-A, ΔIVS-B) Precipitate
- 3: pET-8c Supernatant Liquid
- 4: pET-8c Supernatant Liquid \times 5
- 5: pET-pol (ΔIVS-A, ΔIVS-B) Supernatant Liquid
- 6: pET-pol (ΔIVS-A, ΔIVS-B) Supernatant Liquid \times 5

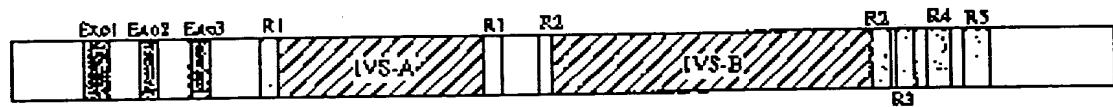
Fig. 6



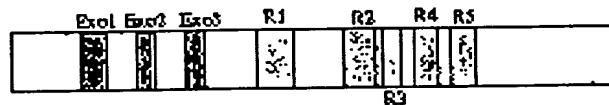
- 1: Vent Polymerase (Derived from *Thermococcus litoralis*)
- 2: pET-pol (ΔIVS-A, ΔIVS-B) Supernatant Liquid
- 3: pET-pol (ΔIVS-A, ΔIVS-B) Supernatant Liquid x 5
- 4: pET-8c Supernatant Liquid
- 5: pET-8c Supernatant Liquid x 5

Fig. 7

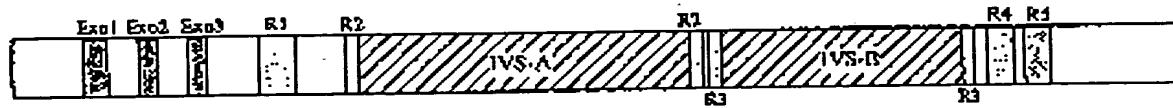
DNA Polymerase Gene of Hyperthermophilic Archaeon Strain K001



DNA Polymerase Gene of Pyrococcus furiosus (Pfu DNA Polymerase)



DNA Polymerase Gene of Thermococcus litoralis (Vent DNA Polymerase)







(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 745 675 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
27.12.1996 Bulletin 1996/52

(51) Int. Cl.⁶: **C12N 15/54, C12N 9/12,
C12N 1/21, C12N 15/70,
C12Q 1/68
// (C12N1/21, C12R1:19)**

(21) Application number: 96108613.9

(22) Date of filing: 30.05.1996

(84) Designated Contracting States:
CH DE FR GB IT LI

(30) Priority: 31.05.1995 JP 134096/95

(71) Applicant: Toyo Boseki Kabushiki Kaisha
Osaka-shi Osaka 530 (JP)

(72) Inventors:

- Kitabayashi, Masao,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)
- Arakawa, Taku,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)
- Inoue, Hiroaki,
c/o Toyo Boseki K.K.
Tsuruga-shi, Fukui-ken (JP)
- Takagi, Masaniro
Suita-shi, Osaka-fu (JP)
- Morikawa, Masaaki
Minoo-shi, Osaka-fu (JP)

(74) Representative: VOSSIUS &
Postfach 86 07 67
81634 München (DE)

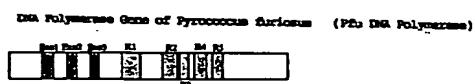
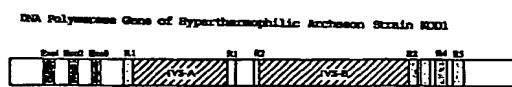
(74) Representative: VOSSIUS & PARTNER
Postfach 86 07 67
81634 München (DE)

(54) A method of amplifying nucleic acid and a reagent therefor

(57) An object of this invention is to provide an enzyme which amplifies nucleic acid with a high fidelity within a short reaction time and also to provide a method of amplification.

This invention relates to a thermostable DNA polymerase having at least 30 bases/second of DNA extension rate and a 3'-5' exonuclease activity derived from an hyperthermophilic archaeon strain KOD1, to a method of amplifying and also of detecting the nucleic acid using said enzyme and to a reagent kit used for those methods.

三





European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 96 10 8613

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
X	DATABASE EMBL DNA-Sequence Data Library, Heidelberg, BRD creation AC D29671, 23 April 1994 XP002017309 * abstract *	1-5, 10, 11	C12N15/54 C12N9/12 C12N1/21 C12N15/70 C12Q1/68 //(C12N1/21, C12R1:19)
Y	& UNPUBLISHED, KAKIHARA H.: "Cloning and analysis of the DNA polymerase gene from a new hyperthermophilic archaeon, Pyrococcus sp. strain KOD1" ---	6-9, 12-20	
X	EP-A-0 624 641 (HOFFMANN-LA ROCHE) 17 November 1994	1,6,10, 15-20	
Y	* the whole document *	6-9, 12-20	
X	EP-A-0 547 920 (NEW ENGLAND BIOLABS, INC.) 5 23 June 1993	5	
A	* the whole document *	1-20	
P,X	FASEB JOURNAL, vol. 10, no. 6, 30 April 1996, page 1423 XP002017308 KITABAYASHI M.: "KOD DNA-Polymerase - Application for fast and accurate PCR" * the whole document *	1-20	C12N C12Q
<p>The present search report has been drawn up for all claims</p>			
Place of search	Date of completion of the search	Examiner	
THE HAGUE	30 October 1996	Kania, T	
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document			